

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:20:35 ; Search time 72.5751 Seconds
(without alignments)
1454.845 Million cell updates/sec

Title: US-10-617-038-1
Perfect score: 1419
Sequence: 1 VEPKSRLLVCAPEPSHARE.....YDNLGLITPTGQGVADGLA 273

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1419	100.0	273	8 ADI37280	Adi37280 M. tuberc
2	121	8.5	1863	8 ADOI3846	Adoi3846 Protein e
3	112.5	7.9	386	7 ABO79432	Abo79432 Pseudomon
4	109.5	7.7	630	7 ABO76880	Abo76880 Pseudomon
5	108.5	7.6	555	8 ADS28926	Ads28926 Bacterial
6	107.5	7.6	717	6 ABU26165	Abu26165 Protein e
7	106	7.5	705	7 ADB74255	Adb74255 Mycobacte
8	104.5	7.4	4083	8 ADOI3847	Adoi3847 Protein e
9	104	7.3	459	7 ABO71375	Abo71375 Pseudomon
10	103.5	7.3	581	8 ADN21896	Adn21896 Bacterial
11	102.5	7.2	807	7 ABO84246	Abo84246 Pseudomon
12	102.5	7.2	931	7 ABO73804	Abo73804 Pseudomon
13	101.5	7.2	250	2 AAY41377	Aay41377 Human sec
14	101.5	7.2	516	8 ADN24653	Adn24653 Bacterial
15	101	7.1	364	6 ABU34030	Abu34030 Protein e
16	100.5	7.1	1024	6 ABU33926	Abu33926 Protein e
17	100.5	7.1	5754	5 ABP62761	Abp62761 S. roseos
18	100.5	7.1	5830	7 ADJ72173	Adj72173 Streptomy
19	100	7.0	321	7 ABO68582	Abo68582 Pseudomon
20	100	7.0	362	7 ABO81477	Abo81477 Pseudomon
21	99.5	7.0	536	7 ABO74749	Abo74749 Pseudomon
22	99	7.0	436	7 ABO71474	Abo71474 Pseudomon
23	98.5	6.9	268	7 ABO82533	Abo82533 Pseudomon
24	98.5	6.9	654	7 ABO73857	Abo73857 Pseudomon
25	98.5	6.9	1267	4 ABG22790	Abg22790 Novel hum

26	98	6.9	220	5 ABP66125	Abp66125 Bifidobac
27	97.5	6.9	897	7 ABO80282	Abo80282 Pseudomon
28	97.5	6.9	3668	7 AD86072	Ad86072 Streptomy
29	97	6.8	350	7 ABO83402	Abo83402 Pseudomon
30	97	6.8	396	7 ABO74580	Abo74580 Pseudomon
31	97	6.8	468	7 ABO70280	Abo70280 Pseudomon
32	97	6.8	499	7 ABO81513	Abo81513 Pseudomon
33	96.5	6.8	1616	8 ADP22958	Adp22958 PRO polyp
34	96.5	6.8	1682	8 ADRI4131	Adri4131 Human NF-
35	96.5	6.8	19938	6 ABB98398	Abb98398 Streptomy
36	96	6.8	307	7 ABO73762	Abo73762 Pseudomon
37	96	6.8	383	6 ABM15872	Abm15872 Mycobacte
38	96	6.8	383	6 ABU36941	Abu36941 Protein e
39	96	6.8	383	6 ABU34879	Abu34879 Protein e
40	96	6.8	408	8 ADS23130	Ads23130 Bacterial
41	96	6.8	472	7 ABO83322	Abo83322 Pseudomon
42	96	6.8	955	7 ABO75508	Abo75508 Pseudomon
43	96	6.8	4630	2 AAW19629	Aaw19629 Streptomy
44	96	6.8	4630	3 AAY77177	Aay77177 S. venezu
45	96	6.8	5215	8 ADU91930	Adu91930 Streptomy

ALIGNMENTS

RESULT 1
ADI37280
ID ADI37280 standard; protein; 273 AA.
XX
AC ADI37280;
XX
DT 22-APR-2004 (first entry)
XX
DE M. tuberculosis low oxygen induced antigen Rv0079 SEQ ID NO:1.
XX
KW mycobacterial infection; vaccine; tuberculosis;
KW Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;
XX low oxygen induced antigen.
XX
OS Mycobacterium tuberculosis.
PN WO2004006952-A2.
XX
PD 22-JAN-2004.
XX
PF 08-JUL-2003; 2003WO-DK000477.
XX
PR 13-JUL-2002; 2002DK-00001098.
XX (STAT-) STATENS SERUM INST.
XX Andersen P, Rosenkrands I, Stryhn A;
WPI: 2004-122778/12.
N-PSDB; ADI37325.
Use of one or more polypeptides or their fragments, which are expressed during the latent stage of the mycobacterial infection, and/or nucleic acids encoding the polypeptides, for a therapeutic vaccine against tuberculosis.
Claim 3; SEQ ID NO 1; 76pp; English.
The present invention describes polypeptides or their fragments, which are expressed during the latent stage of a mycobacterial infection, and/or nucleic acids encoding the polypeptides, which are useful for creating a therapeutic vaccine against tuberculosis. Also described: (1) a therapeutic vaccine against tuberculosis comprising one or more polypeptides; (2) a method for treating an animal, including a human being, with tuberculosis caused by virulent mycobacteria, e.g. by Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for immunising an animal, including a human being, against tuberculosis caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis

CC caused by virulent mycobacteria in an animal, including a human being;
 CC (5) a method for diagnosing previous or ongoing infection with a virulent
 CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
 CC infection in a subject. The polypeptides have antibacterial activities,
 CC and can be used in vaccines and in gene therapy. The polypeptides are
 CC useful for the manufacture of a therapeutic vaccine for treating an
 CC individual who is infected by a virulent mycobacterium, e.g. M.
 CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
 CC The present sequence represents a low oxygen induced antigen, which is
 CC used in the exemplification of the present invention.

XX
 SQ Sequence 273 AA;
 Query Match 100.0%; Score 1419; DB 8; Length 273;
 Best Local Similarity 100.0%; Pred. No. 1.1e-137; Mismatches 0; Indels 0; Gaps 0;
 Matches 273; Conservative 0;
 QY 1 VEPKSRLLVVCAPESHPAREFPDVAVFSGGRANASQAEERLARAVGRVLADRGVTGGARVR 60
 DB 1 VEPKSRLLVVCAPESHPAREFPDVAVFSGGRANASQAEERLARAVGRVLADRGVTGGARVR 60
 QY 61 LTMANCAGDPTLVQINLQVGDTPPLRAQAATAGIDDLRALIRLDRQIVRASAQWCPRPWP 120
 DB 61 LTMANCAGDPTLVQINLQVGDTPPLRAQAATAGIDDLRALIRLDRQIVRASAQWCPRPWP 120
 QY 121 DRPRRLTTPAEALVTRRKPVVLRATPLQATAMDMYDVHLFTDAETGDAVYVRAG 180
 DB 121 DRPRRLTTPAEALVTRRKPVVLRATPLQATAMDMYDVHLFTDAETGDAVYVRAG 180
 QY 181 PSGLRLARQHHVFPFGWSCRAPAGPPVPLIVNSRPTVLTAAAVDRAREHGLPFLPFT 240
 DB 181 PSGLRLARQHHVFPFGWSCRAPAGPPVPLIVNSRPTVLTAAAVDRAREHGLPFLPFT 240
 QY 241 DQATRGQLLYSRDGNLGLITPTGGVADGLA 273
 DB 241 DQATRGQLLYSRDGNLGLITPTGGVADGLA 273

RESULT 2
 ADO13846
 ID ADO13846 standard; protein; 1863 AA.
 XX
 AC ADO13846;
 XX
 DT 15-JUL-2004 (first entry)
 XX
 DE Protein encoded by dbv gene cluster ORF16.
 XX
 KW dbv; Gene cluster; Actinomadura sp.; biosynthesis; glycopeptide; A40926;
 KW 4-hydroxy-phenylglycine; 3; 5-dihydroxy-phenylglycine;
 KW heptapeptide skeleton; chlorination; aromatic residue;
 KW beta-hydroxylation; tyrosine; cross-linking; N-acetyl glucuroamine;
 KW mannoseyl; export; antibiotic.
 XX
 OS Actinomadura sp. ATCC 39727.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1
 FT /note= "Encoded by GTG"

EP1413626-A1.
 PN
 XX 28-APR-2004.
 XX
 PF 23-OCT-2002; 2002EP-00023597.
 XX
 PR 23-OCT-2002; 2002EP-00023597.
 XX
 PA (VICU-) VICURON PHARM INC.
 XX
 PI Donadio S, Sosio M, Beltrametti F;
 XX
 WP1; 2004-332813/31.

DR N-PSDB; ADO13830.
 XX Novel isolated polypeptide A40926, comprising polypeptide sequence or
 PT open reading frame polypeptide involved in biosynthetic pathway of
 PT A409626, useful for synthesizing glycopeptide antibiotic A409626.
 XX
 PS Claim 28; Page 87-95; 164pp; English.
 XX
 CC This sequence is encoded by the dbv gene cluster from *Nonomura* sp.
 CC ATCC39727 (formerly *Actinomadura* sp. ATCC39727). The 37 proteins encoded
 CC by the gene cluster are involved in the biosynthesis of the glycopeptide
 CC A40926. dbv ORF 1, 2, 5, and 37 encode polypeptides required for the
 CC synthesis of 4-hydroxy-phenylglycine residues of A40926. dbv ORF 10-34
 CC and 37 encode polypeptides required for the synthesis of 3, 5-dihydroxy-
 CC phenylglycine residues of A40926. dbv ORF 16, 17, 25, 26 and 36 encode
 CC polypeptides required for the synthesis of heptapeptide skeleton of
 CC A40926. dbv ORF 10 encodes polypeptide required for the chlorination of
 CC the aromatic residues of A40926. dbv ORF 28 encodes a polypeptide
 CC required for the beta-hydroxylation of the tyrosine residue of amino acid
 CC 6 of A40926. dbv ORF 11-14 encode polypeptides required for cross-linking
 CC of the aromatic residues of amino acids at positions 2 and 4, 4 and 6, 1
 CC and 3, and 5 and 7 of A40926. dbv ORF 9, 23 and 29 encode polypeptides
 CC required for addition and formation of N-acetyl glucuroamine residue of
 CC A40926. dbv ORF 20 or 27 encode polypeptides required for the attachment
 CC of mannoseyl residues or N-methylation of A40926. dbv ORF 7, 18, 19, 24
 CC and 35 encode polypeptides required for the export of A40926 or its
 CC precursors. dbv ORF 3, 4, 6, and 22 encode polypeptides required for
 CC regulating the expression of one or more genes of the dbv gene cluster.
 CC The dbv gene cluster and the proteins encoded by it are useful for
 CC producing glycopeptide antibiotic A40926 or its precursor.
 XX
 SQ Sequence 1863 AA;

Query Match 8.5%; Score 121; DB 8; Length 1863;
 Best Local Similarity 27.2%; Pred. No. 0.014;
 Matches 73; Conservative 26; Mismatches 97; Indels 72; Gaps 15;
 QY 7 RLVCAPESHPAREFPDVAVFSGGRANASQAEERLARAVGRVLADRGVTGGARVLTMA 64
 DB 474 RRVATAPDA-----VAITDAGGADLTVAEVDOAANRLARHAAAGIGRDRGVVMD 525
 QY 65 NCADGPTLVQINLQVGDTPPLRAQAATAGIDDLRPA-----LIRLDRQIVRASAQWCPRPWP 120
 DB 526 R---SPDLLIAFL---ASWKAGAAVVPDVEHPAERTFVLADSGV---SAVLCTR--- 572
 QY 121 DRPRRLTTPAEALV-----TRRKPVVLRATPLQATAMDMYDVHLFTDAETG----- 171
 DB 573 ---ATREVAPADAIVIDAPETRAAIDAGAATAPOQIRLSADDLAY--VMYTSGSTGLPKGV 627
 QY 172 ---EDAVV-----YRAGPSGLRLARQHHVFPFG-----WSCRAPAGPPVPLIVNSRPT 217
 DB 628 GVPHGAVAGLAGDEGWRIGPGDVAVLMUATHVFDPSLYAMW-----VPLAMGGR-- 675
 QY 218 PVLTEAAAVD-----RAREHGLPFLPFT 240
 DB 676 VLTPEGVLDALGMQRQAVRGVTFVHLT 703

RESULT 3
 ABO79432
 ID ABO79432 standard; protein; 386 AA.
 XX
 AC ABO79432;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Pseudomonas aeruginosa polypeptide #11607.
 XX
 KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 XX
 OS Pseudomonas aeruginosa.
 XX
 PN US6551795-B1.

XX PD 22-APR-2003.
 XX PF 18-FEB-1999; 99US-00252991.
 XX PR 18-FEB-1998; 98US-0074788P.
 XX PR 27-JUL-1998; 98US-0094190P.
 XX PA (GENO-) GENOME THERAPEUTICS CORP.
 XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 XX DR WPI: 2003-615309/58.
 XX DR N-PSDB; ABD13003.
 XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 XX PT pathological conditions resulting from bacterial infection.
 XX PS Disclosure; SEQ ID NO 28178; 455pp; English.
 XX CC The invention relates to Pseudomonas aeruginosa polypeptides and the
 XX CC polynucleotides encoding them. The sequences are useful in diagnosis and
 XX CC therapy of pathological conditions, as molecular targets for diagnostics,
 XX CC prophylaxis and treatment of pathological conditions resulting from a
 XX CC bacterial infection, for evaluating a compound, such as a polypeptide,
 XX CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 XX CC effective antibacterial targets, as targets for antibacterial drugs,
 XX CC including anti-P. aeruginosa drugs, as templates for recombinant
 XX CC production of P. aeruginosa-derived peptides or polypeptides, as target
 XX CC components for diagnosis and/or treatment of P. aeruginosa-caused
 XX CC infection, and in detection of P. aeruginosa sequences or other sequences
 XX CC of Pseudomonas species using biochip technology. Sequences AB067826-
 XX CC AB084396 represent P. aeruginosa polypeptides of the invention. Note: The
 XX CC sequence data for this patent did not form part of the printed
 XX CC specification but was obtained in electronic format from USPTO at
 XX CC seqdata.uspto.gov/sequence.html
 XX SQ Sequence 386 AA;
 Query Match 7.9%; Score 112.5; DB 7; Length 386;
 Best Local Similarity 24.0%; Pred. No. 0.012;
 Matches 81; Conservative 14; Mismatches 105; Indels 137; Gaps 14;
 QY 1 VEPKSRLLVVC--APPFSHAREFPDVAVFSGGRANASQAERLARAVGRVLRGTGGAR 58
 DB 63 VRPPRRRRCPLPIPGHGR-----RDTGSLCRRRAGPGGR 100
 QY 59 VRLTMANCADGPTLVQINLVQVDTPLR-----AQAATAGIDDL 96
 DB 101 HRQTATRPATGPRHPAAQRPPADRSLSRGLSPRRPPARRQRDGRGRGTARPSATGSPAL 160
 QY 97 R-PALIRLDROIIVRASQWCPWPDPDP-----RRRLTTPA-----EALVT 136
 DB 161 RSPAAGRPARRRARASFGAAGPGRPVPRPARQRLPGRHGRGUSPARQGRQALPALPA 220
 QY 137 RRKPVVLR-----ATPLQIAAMDAMDYDHLFTDAETGEDAVVYRAG 181
 DB 221 RRSFGALPRQLRRLRLRTTARQPATPLSPRGAC-----P 255
 QY 182 SGURL-----ARQHVFPFGWSCRCPAGAP-PVPLIVNSRPTPLTEAA----- 224
 DB 256 TRARLAGNWLPLYNARQPH--REGRSHARSPRLPRLTLPAGPAQPAAGTAAVRPA 313
 QY 225 -----AVD--RAREHGLPELFTFD--QATGRGOL 249
 DB 314 RHFGPALGLDTQLRRRPAGVGLALLFDGQAGAGRL 350
 RESULT 4
 ID AB076880
 XX AB076880 standard; protein; 630 AA.

AC AB076880;
 XX DT 29-JUL-2004 (first entry)
 XX DE Pseudomonas aeruginosa polypeptide #9055.
 XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 XX OS Pseudomonas aeruginosa.
 XX PN US6551795-B1.
 XX PD 22-APR-2003.
 XX PF 18-FEB-1999; 99US-00252991.
 XX PR 18-FEB-1998; 98US-0074788P.
 XX PR 27-JUL-1998; 98US-0094190P.
 XX PA (GENO-) GENOME THERAPEUTICS CORP.
 XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 XX DR WPI: 2003-615309/58.
 XX DR N-PSDB; ABD10451.
 XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 XX PT pathological conditions resulting from bacterial infection.
 XX PS Disclosure; SEQ ID NO 25626; 455pp; English.
 XX CC The invention relates to Pseudomonas aeruginosa polypeptides and the
 XX CC polynucleotides encoding them. The sequences are useful in diagnosis and
 XX CC therapy of pathological conditions, as molecular targets for diagnostics,
 XX CC prophylaxis and treatment of pathological conditions resulting from a
 XX CC bacterial infection, for evaluating a compound, such as a polypeptide,
 XX CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 XX CC effective antibacterial targets, as targets for antibacterial drugs,
 XX CC including anti-P. aeruginosa drugs, as templates for recombinant
 XX CC production of P. aeruginosa-derived peptides or polypeptides, as target
 XX CC components for diagnosis and/or treatment of P. aeruginosa-caused
 XX CC infection, and in detection of P. aeruginosa sequences or other sequences
 XX CC of Pseudomonas species using biochip technology. Sequences AB067826-
 XX CC AB084396 represent P. aeruginosa polypeptides of the invention. Note: The
 XX CC sequence data for this patent did not form part of the printed
 XX CC specification but was obtained in electronic format from USPTO at
 XX CC seqdata.uspto.gov/sequence.html
 XX SQ Sequence 630 AA;
 Query Match 7.7%; Score 109.5; DB 7; Length 630;
 Best Local Similarity 24.6%; Pred. No. 0.049;
 Matches 78; Conservative 15; Mismatches 121; Indels 103; Gaps 12;
 QY 3 PKRSRLWCAPFSSHAREFPDVAVFSGG---RANASQAERLARAVGR---VLADRGV--T 54
 DB 16 PRRRSRRRAAAGRRRR---ARRGGGTTPRRRRHARRRGAGRGSGALADRGVRL 71
 QY 55 GGARV-RLTMANCADG-----PTLVQINLVQVDTPLRAQAATAG----- 92
 DB 72 GPALRHVPVAESQPGTAPTARRPAGRPGRHWAQVORDEVLDRGAGGARGGRL 131
 QY 93 -----IDD-----LRPALIRLDROIIVRASQWCPWPDPDP 124
 DB 132 APCRSPAVGDQAPRRFRGGGDRTRGGQALRPGDSRTRRLAALGPGAPAPPAQOHR 191
 QY 125 R-----RLTTPABALVTRRKVPVLRRLRATPLQAIAMDAMDYDHLFTDAETGEDAVVYRAG 180
 DB 192 RPRSARGGGRRARLAVARRRPAVPATGAPGLPRPAM-----RHGHRPGAPRTG 238
 QY 181 PSGLRLAROHVFPFGW-----SRCPAPAGPPVPLIVNSRPT 218

CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC the gene product or that has an activity against a biological pathway;
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 717 AA;

Query Match 7.6%; Score 107.5; DB 6; Length 717;
 Best Local Similarity 22.7%; Pred. No. 0.093;
 Matches 62; Conservative 29; Mismatches 107; Indels 75; Gaps 11;
 QY 9 VVCAPFESHAREFPDVAVFSGGRANASQAERLARAVGRVLDAGVTGGARVLTMANCAD 68
 DB 322 LLSAPPKQWEQAVDTRDIGALNAGQITELGRTLATPTD-----PRLGASLHCGG 375
 QY 69 G--PTLVQINLVQV-GDTPLRAQAATAGIDDLRLPALRLDROIVRAS-----AOWC 115
 DB 376 GAAPTIANLSLDARGDIP-RAIAMS HQ-----RFTREVARLAGFVDPDSECTGEAI 427
 QY 116 PRWPDRPRRR-----LTTPAELVTRKPKPVLLRRATPLQA 151
 DB 428 ARAMPKNVARRDGDVSLMAGSRAELGVSLSGHEWLAQAQSLTHRGSVI-RSAAPISE 486
 QY 152 IAAWDAMDYDVHLFTAEETGEDAVVVRAGPSGLRLARQHVVFPFGNSCRAPAGPVPLI 211
 DB 487 AAADVIGV-----AERTEATLI-----DGSURGTIRH-----AGAITLST 523
 QY 212 VNSRPTPVLTAAAVDRAREHGLPFLFFFTDQAT 244
 DB 524 TNVKEGELAEQALAGIRTHGLELFTFSKAT 556

RESULT 7
 ADB74255
 ID ADB74255 standard; protein; 705 AA.
 XX
 AC ADB74255;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Mycobacterium tuberculosis non-naturally occurring peptide #3.
 XX
 KW Non-naturally occurring peptide; anion pump protein; tuberculosis;
 KW hypersensitivity reaction; tuberculostatic.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN US6583266-B1.
 XX

PD 24-JUN-2003.
 XX
 PF 16-SEP-1994; 94US-00311731.
 XX
 PR 19-AUG-1993; 93US-00109181.
 PR 22-OCT-1993; 93US-00142558.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 PI Smith DR, Mao J;
 XX WPI; 2003-556441/52.
 DR N-PSDB; ADB74252.
 XX
 PT New Mycobacterium tuberculosis anion pump peptide useful for as
 XX tuberculosis vaccine and diagnosis of tuberculosis infection.
 PS Disclosure; SEQ ID NO 4; 26pp; English.
 XX
 CC The invention relates to a non-naturally occurring peptide of
 CC Mycobacterium tuberculosis comprising an amino acid sequence
 CC corresponding to an anion pump protein. The invention also relates to a
 CC non-naturally occurring nucleic acid corresponding to a DNA sequence of
 CC Mycobacterium tuberculosis or Mycobacterium leprae. The new peptide is
 CC useful as a vaccine against Mycobacterium tuberculosis or Mycobacterium
 CC leprae or for screening for new tuberculosis drugs. Purified proteins
 CC derived from the sequences of the invention may elicit a specific immune
 CC response. The peptide may also be used to detect hypersensitivity
 CC reactions of individuals exposed to Mycobacterium tuberculosis or
 CC Mycobacterium leprae. The proteins and peptides may be affixed to solid
 CC supports to detect antibodies typical of hypersensitivity reactions, from
 CC a patient's sera. This sequence represents Mycobacterium tuberculosis non
 CC -naturally occurring peptide of the invention. Note: The sequence data
 CC for this patent did not form part of the printed specification but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 705 AA;

Query Match 7.5%; Score 106; DB 7; Length 705;
 Best Local Similarity 24.6%; Pred. No. 0.13;
 Matches 65; Conservative 31; Mismatches 104; Indels 64; Gaps 10;
 QY 12 APEPESHAREFPDVAVFSGGRANASQAERLARAVGRVLDAGVTGGARVLTMANCADGPT 71
 DB 11 AEQASEAGWYDRPAFYAADVVTHGQIHDCGAEVLNRGLSSGDRVLLCL---PDSPD 67
 QY 72 LVQINLVQVGDTPRLRAQAATAGIDDLRLPALRLDROIVRASQAOWCPRPWDRPRRLTTPA 131
 DB 68 LVQILLACLARGVMAFLA-----NPELHRDDHALARNT-----PALVVT 109
 QY 132 EALVTRKPKPVLLRRATPLQAIAM-----DAMDY-----DVHLFTD 167
 DB 110 DALRDRFQPSRVAEAAELMSEARVAPGYPGEMGGDALAYATYTTGTPPKAAIHRHAD 169
 QY 168 AETGEDAVVYRA-----GPSGLRLARQHVVFPFG---WSRCRAPAGPVPLIVNSRPTPV 219
 DB 170 PLTFVDAMCRKALRUTPEDTGLCSARMYFAYGLGNSVW-----FPLATGGSVINS--APV 223
 QY 220 LTEAAAVDRAR-----EHLPLPLF 238
 DB 224 TPEAAAILSARFGPSVLYGVNPF 247

RESULT 8
 ADO13847
 ID ADO13847 standard; protein; 4083 AA.
 XX
 AC ADO13847;
 XX
 DT 15-JUL-2004 (first entry)
 XX
 DE Protein encoded by dbv gene cluster ORF17.

XX dbv: gene cluster; Actinomadura sp.; biosynthesis; glycopeptide; A40926;
KW 4-hydroxy-phenylglycine; 3; 5-dihydroxy-phenylglycine;
KW heptapeptide skeleton; chlorination; aromatic residue;
KW beta-hydroxylation; tyrosine; cross-linking; N-acetyl glucuroamine;
KW mannose; export; antibiotic.
XX Actinomadura sp. ATCC 39727.
XX EP1413626-A1.
XX 28-APR-2004.
XX 23-OCT-2002; 2002EP-00023597.
XX 23-OCT-2002; 2002EP-00023597.
XX (VICU-) VICURON PHARM INC.
XX Donadio S, Sosio M, Beltrametti F;
XX WPI; 2004-332813/31.
XX N-PSDB; ADO13830.
XX Novel isolated polypeptide A40926, comprising polypeptide sequence or
PT open reading frame polypeptide involved in biosynthetic pathway of
PT A40926, useful for synthesizing glycopeptide antibiotic A40926.
XX Claim 28; Page 95-112; 164pp; English.
XX This sequence is encoded by the dbv gene cluster from Nonomura sp.
CC ATCC39727 (formerly Actinomadura sp. ATCC39727). The 37 proteins encoded
CC by the gene cluster are involved in the biosynthesis of the glycopeptide
CC A40926. dbv ORF 1, 2, 5, and 37 encode polypeptides required for the
CC synthesis of 4-hydroxy-phenylglycine residues of A40926. dbv ORF 30-34
CC and 37 encode polypeptides required for the synthesis of 3, 5-dihydroxy-
CC phenylglycine residues of A40926. dbv ORF 16, 17, 25, 26 and 36 encode
CC polypeptides required for the synthesis of heptapeptide skeleton of
CC A40926. dbv ORF 10 encodes polypeptide required for the chlorination of
CC the aromatic residues of A40926. dbv ORF 28 encodes a polypeptide
CC required for the beta-hydroxylation of the tyrosine residue of amino acid
CC 6 of A40926. dbv ORF 11-14 encode polypeptides required for cross-linking
CC of the aromatic residues of amino acids at positions 2 and 4, 4 and 6, 1
CC and 3, and 5 and 7 of A40926. dbv ORF 9, 23 and 29 encode polypeptides
CC required for addition and formation of N-acetyl glucuroamine residue of
CC A40926. dbv ORF 20 or 27 encode polypeptides required for the attachment
CC of mannose residues or N-methylation of A40926. dbv ORF 7, 18, 19, 24
CC and 35 encode polypeptides required for the export of A40926 or its
CC precursors. dbv ORF 3, 4, 6, and 22 encode polypeptides required for
CC regulating the expression of one or more genes of the dbv gene cluster.
CC The dbv gene cluster and the proteins encoded by it are useful for
CC producing glycopeptide antibiotic A40926 or its precursor.
XX Sequence 4083 AA;
Query Match 7.4%; Score 104.5; DB 8; Length 4083;
Best Local Similarity 26.8%; Pred. No. 2;
Matches 67; Conservative 21; Mismatches 69; Indels 93; Gaps 15;
QY 24 VAVFSGGRR-----NASQERLARAVGRVLADRGVTGGARVLTWANCADGPTLVQINLQ 78
DB 472 VAVSEGKRAVSGELAEHAERLA---GYLSGRGVRRGRVAVVMGR---SGLIATLLA 524
QY 79 VCDTFLRAQAATAGTDDLRLPALIRLDR-QIVRASQAQWCPRPDRPRRLTTTPAEALVTR 137
DB 525 V-----WKAGAAFPVPDPAVPA-----ERVQFMADAE-----PA-AVYTE 559
QY 138 R-----KPVVLRATPLQIAAMDAMDYDVH-----LFTDAETGEDAVVYRA 179
DB 560 RACQAAVPAAGGLDPIVLDPPDLRAVAEHARLSAGAHADDLAYVNYTSGTGR----- 612
QY 180 GPSGLRLARQHHVF---PPGWSRCRAPAGPP-----VPLVNSRPT 217
| : : : : : |

DB 613 -PKGVAVSHGNVAALAGEPGWG-----LGPEDAVLHMAHAFDISLFELWVPLLSGAR-- 664
QY 218 PVLTEAAAVD 227
DB 665 VVLAEPGAVD 674
RESULT 9
ABO71375
ID ABO71375 standard; protein: 459 AA.
XX ABO71375;
XX ABO71375;
XX 29-JUL-2004 (first entry)
XX Pseudomonas aeruginosa polypeptide #3550.
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX Pseudomonas aeruginosa.
XX US6551795-B1.
XX 22-APR-2003.
XX 18-FEB-1999; 99US-00252991.
XX 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
WPI; 2003-615309/58.
XX N-PSDB; ABD04946.
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX Disclosure; SEQ ID NO 20121; 455pp; English.
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX Sequence 459 AA;
Query Match 7.3%; Score 104; DB 7; Length 459;
Best Local Similarity 23.4%; Pred. No. 0.12;
Matches 77; Conservative 19; Mismatches 109; Indels 124; Gaps 12;
QY 3 PKRSLVVCAPESPHAREFPDVAVFSGGRANASQAERLARAVGRVLADRGVTGG----- 56
DB 2 PPGARVAGLRPPAGARRP-VAPVRRARPGARRLEQPARVGV-VAXARGLPGGRLRAA 59
QY 57 -----ARVRLTWANCADGPTLVQINLQVGDPTFLRAQAATAGI----- 93
DB 60 RWPFRAREVNDPHGETRRERLCATGLAAGFVRLRRRAALG-----RURARPGRGAGH 115
| : : : : : |

QY 94 -DLIRPALI-----RDDRQIVRASAQW 114
D 116 GDLRPAQVGAASAGRGQAHPARRPACGGRAADPHGTDPGGQRRFAAQVQVORA-- 173
QY 115 CRRPWPDRPRRLTTAE-----ALVTRKPVVLRATPLQAIAAMDMDYDV 162
D 174 ---AQGAPAGRTRPADPGDARGAGRAGPAPDRAPGV-----AAVAAQPPD- 220
QY 163 HLFTDAETGDAVYVYRAGPSGLRLARQHVVFPFGWSCRAPAGPPVPLVINSRPTV--- 219
D 221 -----RAGQRTLAIGEHRGAARPARVAPDRGQAP-----ATTVEQP 260
QY 220 LTEAAAVDRAREHGLPFLFTDQATGRGQ 248
D 261 VERCARPRRGRHLHAAQPVARTGPTGRGE 289

RESULT 10
ADN21896

ID ADN21896 standard; protein; 581 AA.

AC ADN21896;

DT 02-DEC-2004 (first entry)

XX Bacterial polypeptide #4549.

XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

OS Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

XX (HINK/) HINKLE G J.

XX (SLAT/) SLATER S C.

XX (CHEN/) CHEN X.

XX (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide

XX for expression of a polynucleotide encoding a polypeptide from a

XX microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 4549; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with

CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 581 AA;

Query Match 7.3%; Score 103.5; DB 8; Length 581;

Best Local Similarity 24.0%; Pred. No. 0.18;

Matches 81; Conservative 26; Mismatches 100; Indels 131; Gaps 18;

QY 13 PEP-----SHAREFPDVAVFSGG--RANASQAERLARAVGRVLADRGVTGG 56

D 38 PEPDRYPAMRNLDLIQSAARTPDAESLVCGDVRITYQLQSSRAFGQVLSLGVAPG 97

QY 57 ARVRL-----TMANCADGPTLVQIN-----LQVG-----DTPURAQA 88

D 98 GRVAIFLDRKMETVVMGLGACAGCVFVPVNPLLKPEQVAHVLRDSCARCLVTTALRVRA 157

QY 89 ----ATAGIDDI-----RPALIRLQRIVRASAQWCPRPWPDRPRRLTTPAE 132

D 158 LAEGVAPVTDILVDEPQVTPDQPTAVRVHR-----WPD-----CISSPDD 200

QY 133 ALVTRRKPVLRARRATPLQAIAMDMDYDVHLFTDAETG-----EDAV 175

D 201 ALV-----PAPAVTCIDT-DLAALYTSGSTGLAKGWSLHRNLLLEGANSVAE 247

QY 176 VYRAGPSGLRLARQHVVFPFGWSCRCA--PAGEPPVPLVNSRPTVPLTAAAVDRAREH- 232

D 248 YLRLDPSDRVLAVPLSLFDAGLSQLTSAWSAGATAVLL-----NYLTAHDAVLACEREHI 302

QY 233 ----GLPELFFTTDAQTG-----RGOLLYSRVDGNLG 259

D 303 TAITGVPPPLMM--QUTGATWSDAARNTL---RYFANTG 335

RESULT 11

ABO84246

ID ABO84246 standard; protein; 807 AA.

XX ABO84246;

XX 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #16421.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US6551795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

KW inflammation; ischemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.
XX
XX WO9947540-A1.
XX
XX 23-SEP-1999.
XX
XX 18-MAR-1999; 99WO-US005804.
XX
PR 19-MAR-1998; 98US-0078563P.
PR 19-MAR-1998; 98US-0078566P.
PR 19-MAR-1998; 98US-0078573P.
PR 19-MAR-1998; 98US-0078574P.
PR 19-MAR-1998; 98US-0078576P.
PR 19-MAR-1998; 98US-0078577P.
PR 19-MAR-1998; 98US-0078578P.
PR 19-MAR-1998; 98US-0078579P.
PR 19-MAR-1998; 98US-0078581P.
PR 01-APR-1998; 98US-0080312P.
PR 01-APR-1998; 98US-0080313P.
PR 01-APR-1998; 98US-0080314P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Ni J, Rosen CA, Yu G, Young PE, Feng P, Soppet DR;
PI Wei Y, Endress GA, Duan RD, Kyaw H, Ebner R, Lafleur DW, Olsen HS;
PI Shi Y, Moore PA;
XX
XX WPI; 1999-562050/47.
DR N-PSDB; AAZ24880.
XX
XX New isolated human genes, useful for diagnosis and treatment of e.g.
PT cancers, neurological disorders, immune diseases, inflammation or blood
PT disorders.
XX
XX Claim 11; Page 397-398; 484pp; English.
XX
XX This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. AAZ24802) for increasing the stability of the fused protein
CC as compared to the human protein only. The invention relates to 95 novel
CC genes and their fragments (nucleic acid sequences: AAZ24811-224907; amino
CC acid sequences AAY41308-Y41404) which are useful for preventing, treating
CC or ameliorating medical conditions e.g. by protein or gene therapy. Also,
CC pathological conditions can be diagnosed by determining the amount of the
CC new polypeptides in a sample or by determining the presence of mutations
CC in the new polynucleotides. Specific uses are described for each of the
CC 95 polynucleotides, based on which tissues they are most highly expressed
CC in (see AAZ24811 for described uses)
XX
SQ Sequence 250 AA;

Query Match 7.2%; Score 101.5; DB 2; Length 250;
Best Local Similarity 26.0%; Pred. No. 0.093;
Matches 58; Conservative 13; Mismatches 85; Indels 67; Gaps 8;

QY 13 PEPHAREFPDVAVFGSGGRANASQAEERLARAVGRVLADRGVTCGARVRLTMANCADGPTL 72
DB 68 PEPAPAR-----AAAGPQPPCAARCLRGSRATA---PGPARERAAAGACA---SL 114
QY 73 VQNLQVGDTPRLAQATAGIDDLRLPALIRLDRIQIVRASQWCP--RPWDRPRRLTTP 130
DB 115 LGFCRAAAGPERQPGAGSTGTRDFR-AAARAAGQLIGRQPGAGARGARRAPAAALTQP 173
QY 131 AEALVTRKPKWLRRATPLQIAAMDAMDYVHLFTDAETGEDAVVYRAGPSGLRLARQH 190
DB 174 A-----GQRAAGTGAAG----- 187

QY 191 HVFPFGWRCRAPAGPPVPLIVNSRPTVLTAAAVDRAREHG 233
DB 188 ---PPARSRRAPARQPLGLRV--RAAPALRLAAPAPARVRG 225

RESULT 14
ADN24653
ID ADN24653 standard; protein; 516 AA.
XX
XX ADN24653;
XX
XX 02-DEC-2004 (first entry)
XX
XX Bacterial polypeptide #7306.
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
XX Bacteria.
XX
XX US2003233675-A1.
XX
XX 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 7306; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

```

XX SQ Sequence 516 AA;
Query Match
Best Local Similarity 7.2%; Score 101.5; DB 8; Length 516;
Matches 78; Conservative 26; Mismatches 100; Indels 117; Gaps 17;

QY 16 SHAREPDAVPSGG--RANASQERLARAVGRVLADRGVTCGARVRL-----T 62
DB 8 SAARTPDASLVCGDVRITYRQLQSSRAFGVLSIGVAPGGRVAIFLDKRMETVYSM 67
QY 63 MANCAGDPTLVQIN-----LQVG-----DTPLRAQA-----ATAGIDDL----- 96
DB 68 LGACAAGCVFVPNLLKPEQVAHVLRDVGARCLVTTALRVPALAEGVAPVTDLILVDE 127
QY 97 -----RPALIRLDRQIVRASAQCPWPDPRRRLTTTPAEALVTRKPKVVLRRATPL 149
DB 128 PQQVTPDQPTAVRVHR-----WPD-----CISSPDALV-----PA 158
QY 150 QAIAMDAMDYDVHLFTDAETG-----EDAVVYRAGPSGLRLARQHHV 192
DB 159 PAVTCIDT-DLAALTYTSGSTGLAKGVMLSHRNLLLEGASVAEYLRLDPSORVLAVLPLS 217
QY 193 FPGWSCRCA--PAGPPVPLVNSRPTVPLTEAAAVDRAREH-----GLPFLFTDQATG 245
DB 218 FDAGLSQLTSAWSAGATAVLL-----NYLTAHDAVLACEREHITAITGVPLLM--QLTG 270
QY 246 -----RGQLLYRYDGNL 259
DB 271 ATWSDAARNTL---RYFANTG 288

RESULT 15
ABU34030
ID ABU34030 standard; protein; 364 AA.
XX AC ABU34030;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #19557.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Mycobacterium avium.
XX PN WO20027183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX DR N-PSDB; ACA37900.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 61954; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of

```

```

CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC product's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: the sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 364 AA;

```

```

Query Match      7.1%; Score 101; DB 6; Length 364;
Best Local Similarity 26.7%; Pred. No. 0.17;
Matches 82; Conservative 30; Mismatches 89; Indels 106; Gaps 20;

QY 17 HAREFPDVA---VFSGGRANASQ--AERLARAVGR--VLADR---GVTGG--ARVRLTMA 64
DB 90 HQRVIDNVATVPVLRGQRRRAARTAAQVLERGLDKLARRYPAQUSGGQQRGVARA 149
QY 65 NCADGPTLVQINLQVGDTPRAQAATAGID-----DLRPAIRLRDRQIVR----- 109
DB 150 LAADPPILLM-----DEPESA-----VDPVVRHDLQNEILRLQSELHKITVFTHDID 197
QY 110 -----ASQWCPRPWDPDRRLTTTPAEALVTRKPKVVLRRATP-LOATAA 154
DB 198 EALRLGERVAVFKGMLQCC-----DEPARLLSRPANDFVSRF--IGLGRGYRWLQLIDA 250
QY 155 MDANDYDV-HLFTDAETGEDAVVYRAGPSGLRLARQHHVPPGWSRCRAPAGPPVPLIVN 213
DB 251 AGLPVHVRRLFAENLSGAD-----VF-DGAVVVDEAGAPLGI-- 289
QY 214 SRTPVLTEAAAVDRAREHGLPFLFTDQATGRLQLLYSYRDGNL-----GLITPTGDG 267
DB 290 -----DABGLARHRDG-----LALSDSMTGVLGSLFRPR--GNLSQALDAALSSPSGIG 335
QY 268 VA---DG 271
DB 336 VAVDDDG 342

```

Search completed: October 15, 2005, 04:26:14
Job time : 74.5751 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 01:42:44 ; Search time 22.4145 Seconds
(without alignments)
909.197 Million cell updates/sec

Title: US-10-617-038-1
Perfect score: 1419
Sequence: 1 VEPKSRRLVVCAPPEPSSHARE.....YDNLGLITPTGCGVADGLA 273

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Issued Patents AA:*
- 1: /cgn2_6/protdata/1/iaa/5A_COMB.pep.*
 - 2: /cgn2_6/protdata/1/iaa/5B_COMB.pep.*
 - 3: /cgn2_6/protdata/1/iaa/6A_COMB.pep.*
 - 4: /cgn2_6/protdata/1/iaa/6B_COMB.pep.*
 - 5: /cgn2_6/protdata/1/iaa/PCtus COMB.pep.*
 - 6: /cgn2_6/protdata/1/iaa/backfilees1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	112.5	7.9	386	4	US-09-252-991A-28178 Sequence 28178, A
2	109.5	7.7	630	4	US-09-252-991A-25626 Sequence 25626, A
3	106	7.5	705	4	US-08-311-731A-4 Sequence 4, Appli
4	104	7.3	459	4	US-09-252-991A-20121 Sequence 20121, A
5	102.5	7.2	807	4	US-09-252-991A-32992 Sequence 32992, A
6	102.5	7.2	931	4	US-09-252-991A-22550 Sequence 22550, A
7	100	7.0	321	4	US-09-252-991A-17328 Sequence 17328, A
8	100	7.0	362	4	US-09-252-991A-30223 Sequence 30223, A
9	99.5	7.0	536	4	US-09-252-991A-23495 Sequence 23495, A
10	99	7.0	436	4	US-09-252-991A-20220 Sequence 20220, A
11	98.5	6.9	268	4	US-09-252-991A-31279 Sequence 31279, A
12	98.5	6.9	654	4	US-09-252-991A-19026 Sequence 19026, A
13	98	6.9	296	4	US-09-252-991A-13964 Sequence 13964, A
14	97.5	6.9	897	4	US-09-252-991A-23028 Sequence 23028, A
15	97	6.8	350	4	US-09-252-991A-32148 Sequence 32148, A
16	97	6.8	396	4	US-09-252-991A-23326 Sequence 23326, A
17	97	6.8	468	4	US-09-252-991A-19026 Sequence 19026, A
18	97	6.8	499	4	US-09-252-991A-30259 Sequence 30259, A
19	96	6.8	307	4	US-09-252-991A-22508 Sequence 22508, A
20	96	6.8	472	4	US-09-252-991A-32068 Sequence 32068, A
21	96	6.8	955	4	US-09-252-991A-24254 Sequence 24254, A
22	96	6.8	4630	4	US-09-091-609-2 Sequence 2, Appli
23	96	6.8	5215	3	US-09-105-537-2 Sequence 2, Appli
24	95.5	6.7	229	4	US-09-252-991A-27802 Sequence 27802, A
25	95.5	6.7	339	4	US-09-252-991A-32978 Sequence 32978, A
26	95.5	6.7	343	4	US-09-252-991A-28970 Sequence 28970, A
27	95.5	6.7	510	4	US-09-252-991A-25076 Sequence 25076, A

ALIGNMENTS

RESULT 1
US-09-252-991A-28178
; Sequence 28178, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28178
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28178

Query Match		7.9%;	Score 112.5;	DB 4;	Length 386;
Best Local Similarity		24.0%;	Pred. No. 0.00075;		
Matches		81;	Conservative 14;	Mismatches 105;	Indels 137; Gaps 14;
QY	1	VEPKSRRLVVC--APEPSSHAREFPDVAVFSGGRANASQAERLARAVGRVLADRGVTGGAR	58		
DB	63	VRPPRRRRPCRLPIPGHGR-----RDTGSLCRRRAGPGGR	100		
QY	59	VRLTMANACAGTFLVQINLQVGTPLR-----AQAATAGIDDL	96		
DB	101	HRQTATRPATGPRHPAAQRRPADRSLLRGSLLRPPPPARRQDRGRGRRTARPSATGSPAL	160		
QY	97	R-PALIRLDQIVRASAOQCPRPWDRP-----RRRLTTPA-----EALVT	136		
DB	161	RSPAAGRRARRARASPGAPGPRVPRPARQPLRGRHRLGSLPARQFGRQALPRALPA	220		
QY	137	RKKPVVLR-----ARQHHVFPGHSCRCRAPGP-PVPLIVNSRPTPLTEA-----	224		
DB	221	RRSPGALPQLRQRLRLRTTARQATPLSPRGAC-----P	255		
QY	182	SGRL-----ARQHHVFPGHSCRCRAPGP-PVPLIVNSRPTPLTEA-----	224		
DB	256	TRARLAGNWRLLPVNAQPH--RPGHSHARSPLRPLPLTLAGPQAQPAAGTAARVAPGA	313		
QY	225	-----AVD-----RAREHGLPFLFFTD--QATGRGOL	249		
DB	314	RHPGPALGDLTOLRRRPPAGVGLALLPFGGQAQAGRL	350		

D**b** 665 AFASRAPRSLH-----PTGYGTLGYGLPAGIGAKLGAPQRPGLVLVGDGGLYTAQ 716

US-09-252-991A-17328

US-09-252-991A-17328

Query Match 7.0%; Score 100; DB 4; Length 321;
Best Local Similarity 30.2%; Pred. No. 0.014; 94; Indels 60; Gaps 13;
Matches 71; Conservative 10; Mismatches 10; Indels 60; Gaps 13;
QY 28 SGRNASQAERLARAVGRVLADRGVTGARGVRLTMANCADGPTLVQINLVQVGTGTPPRAQ 87
DB 24 AGGPAASRAE--ARLHLRAL-DPLGEGPRHR---CGAAGPVLRL-----PHRPQ 69
QY 88 -AATAGIDDLRAL-----IRLDRQIVRASAQWCP-----RWPDR----- 122
DB 70 RAAAAAPHRIEPTPAQRORRLQALHGRLLPPLPPADPVLPLGLGRGAAGDQPRDRRLLV 129
QY 123 --PRRLA---TTPAEALVTRKPVVL-----RRATPLQAIAMDMYDVHLFTDAETG 171
DB 130 GAPRRALARHAEPCAGARAAGRYRILLHRPPRADPLPGRPADPGR-----VFRGPGG 185
QY 172 ED-----AVVYRAGPSGLRLARQHVFPFGWSRCRAPAG-PPVPLIVNSRPTPVL 220
DB 186 RGIRPMAGAAQAQPPRRRLDEPGHGLPGLGPRPQAQVPPAPCAAGPRISPAL 240
RESULT 8
US-09-252-991A-30223
; Sequence 30223, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30223
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30223

Query Match 7.0%; Score 100; DB 4; Length 362;
Best Local Similarity 26.6%; Pred. No. 0.016; 89; Indels 76; Gaps 14;
Matches 67; Conservative 20; Mismatches 20; Indels 76; Gaps 14;
QY 5 RSRLVVCAPESHPAREFFPDVAVFSGGRA---NASQAERLARAVGRVLADRGVTGARGVRL 61
DB 73 RQAGLCRRDRHARRRATAVGILGGRADPCTGATVRLAE-GRIVA--AVQGPSARE 129
QY 62 TWANCADG--PTLV-----QINLVQVGTGTPRAQAATAGIDDLRALIRLDRQIVRASA 112
DB 130 PAQGAAGTAPAVVPRGAARRVAGRLRRTGLRVALHGA---LRPLAVDVFRQPVAGLV 185
QY 113 QWCP-RPW-----PDRPRRLTTPAE-----ALVTRKPV--V 142
DB 186 GVCPCRPGLYLKRRVFRRLPRLSPARRRRRLSPVRLPALRPRRAAGGTARRPAGRT 245
QY 143 LR-----RATPLQAIAMDMYDVHLFTDAETGDAVVYRAGPSGLRLARQHVFPFGW 197
DB 246 LRQSLAGRAAGAAVVPVRL-----RKAAGFRPGPAL-----SPIQPPR- 286
QY 198 SRCRAPAGFPVP 209
DB 287 ---RAPAGDSQP 295
RESULT 9
US-09-252-991A-23495
; Sequence 23495, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23495
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23495

Query Match 7.0%; Score 99.5; DB 4; Length 536;
Best Local Similarity 26.6%; Pred. No. 0.032;
Matches 69; Conservative 15; Mismatches 108; Indels 67; Gaps 11;
QY 3 PKRSRLVVCAPESHPAREFFPDVAVFSGGRANASQAERLARAVGRVLADRGVTGARGVRLT 62
DB 147 PRRHLLVERRGRPGNGR-----AGGTARGEQLQALHGLQERHGRRRYPGG----- 192
QY 63 MANCADGPTLVQINLVQVGTGTPLRA-QAATAGIDDLRALIRLDRQIVRASAQWCPRPWPD 121
DB 193 -----QRLALPGAGRGADGARG-----ERRTGLPPAAETPRPGPH 227
QY 122 RPRRLTTPAEALVTR---RKPVLVLRATPLQAIAMDMYDVHLFTDAETGDAVVY 177
DB 228 RPRGASSVASAASRCGRQRPHP--HRRDAGYAAVPGA-----YFQPRGAGRDRI-CP 277
QY 178 RAGPSGLR-----LARQHVFPFGWS-RCRAPAGPPVPLIVNSRPTPVLTEA-AAVDR 228
DB 278 RQGPAGLRGAGRAPAAARRQLRPPGLGHRGIRDEPAVP---SRTPGSAVARPAVRQ 333
QY 229 AREHGLPFLFTDOATGRG 247
DB 334 PAYHRHRLLLLRRAEGHG 352

RESULT 10
US-09-252-991A-20220
; Sequence 20220, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20220
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20220

Query Match 7.0%; Score 99; DB 4; Length 436;
Best Local Similarity 27.5%; Pred. No. 0.027;
Matches 76; Conservative 18; Mismatches 114; Indels 68; Gaps 16;
QY 29 GGRANASQAERLARAVGRVLADRGVTGARGVRLTMANCADGPTLVQINLVQVGTGTPRAQA 88
DB 39 GGSRRAGPGGLGDLAQRRKPPATPTGAPELCAGTSLSGGA-----AGGRTEERQRP 92

89	Qy	ATAG--IDDLR-PALIRLDRQIVASAQWCPRPWDPDRPRR---LTTPASAL-----VTRR 138
93	Db	SVAGATLEDRLRPP--RRQQPGRHQLRQPALPGAERPARRQLARLPAAAAAASGLSRQ 150
139	Qy	KPVVLRRATPLQAIAMDAMDYDVHLFTDASTGDAVVYRAGPSGLFLAROHVFPFGWS 198
151	Db	HGQIPRR--TPGRAFRQL-----TGGPDGNNPPSVRP--GLRCLQR---PPALS 193
199	Qy	RCRAP----AGPPVPPLIV-----NSRTPPVLTEA-----AADVDRARE--HGLPFL 237
194	Db	RCALPPSRRLAPAPARLGLRHRYPPGRAPCAVLRRSPQRHQRAADRTTCPCSGYPL- 252
238	Qy	FTDQATGRGQLLYSKYDGNLGLIPTTGDGVADGLA 273
253	Db	----ORAGRRSHRLRRRLRPLGRSP-----GLA 277

```

RESULT 11
US-09-252-991A-31279
; Sequence 31279, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31279
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31279

```

Query Match	6.9%;	Score 98.5;	DB 4;	Length 268;
Best Local Similarity	25.0%;	Pred. No. 0.016;		
Matches 53;	Conservative 11;	Mismatches 69;	Indels 79;	Gaps 7;
Qy	25	AVFSGGGRANASOABERLARAVGRVLADRGVTGGARVRLTMANCADGPTLVQINLQVGTPTL	84	
Db	71	AAASGPGAGRAPRANSAGYRPGLPFARGRASGRPRSPG-----VRGLSAGRSPL	122	
Qy	85	RAQATAGATIDDLRPALIRLQIVRSAQMCPRPWPDRPRRLTTPBAALVTRRKPVVLR	144	
Db	123	RRR-----RAGCRRRRRAGPGRWRRRRAVAPAGCL-----	151	
Qy	145	RATPLQAIAMDAMDYDVLHPTDAETGEDAVVYFAGPSGLRLARQHHVFPFGWSCRAP-	203	
Db	152	RSGP-----VARGGPA-----ARSL-----FPPGAPGCRCPA	177	
Qy	204	-----AGPPVPLIVNSRPTPVLTEAAAVDRAR	230	
Db	178	PPGAGAGCGSPAAASLSPARSVRRAGAGSSAR	209	

RESULT 12
US-09-252-991A-22603
; Sequence 22603, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18

```

; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22603
; LENGTH: 654
; TYPE: prt
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22603

```

Query Match	6.9%;	Score 98.5;	DB 4;	Length 654;		
Best Local Similarity	27.6%;	Pred. NO. 0.055;				
Matches	70;	Conservative	19;	Mismatches 98; Indels 67; Gaps 12;		
Qy	5	RSRLVVCAPPSHAREFPD----	VAVFSGGRANASQ--	AERLARAVGRVLAD-----	50	
Db	15	RPRVALRYPQERHPAPADAGEAQVLYAGGQVPPRCAADPPRQAGRVPADAGRIQPG	74			
Qy	51	--RGVGTGARVRLTMANCADGPTLVQINLQVGDTPLRQAQATAGIDDLRALLRLDQRQIV	108			
Db	75	GLRGVLRGRHRLUA--AGRKLPGLCV-----	RVL	103		
Qy	109	RASAQMCPRPWPDPRRRLTTPAEALVTRKRPVL--	RRATPLQAI	AAMDAMDYDVHLFT	156	
Db	104	RARSQ--LRRPGSTAPLQR--	PATDLVAARWPGHLP	PELRHPPSRA-----	DRDAVGGT	152
Qy	167	DAETGDEAVYRAGPSGLRLARQHHV--	FP--	PGHSRCRAPAGPVPPLVNSRPTPVL	221	
Db	153	DLHQEPAPTAAPGRVYLRRLRYRAGVDRPQYHPGDAELRAAAEPRFPLETDOOHVAQLP	212			
Qy	222	EAAAVDRAREHGLP	235			
Db	213	VAG--QRRRTQGD	224			

```

RESULT 13
US-09-902-540-13964
; Sequence 13964, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13964
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13964

```

Query Match	6.9%	Score 98;	DB 4;	Length 296;
Best Local Similarity	25.4%;	Pred. No. 0.021;		
Matches	60;	Conservative 31;	Mismatches 83;	Indels 62; Gaps 12;
Qy	28	SGGRANASQAEIRIARAVGRVLADRGVTGGARVLTWANCADGPTLQAQINLOVGDPTPLRAQ	87	
Db	5	SVGPVEALLAQLAPAV--VATVHWIPGGALLRVLEAGA--GPTVLLHGKG-----	53	
Qy	88	AATAGIDDLRPALIRLDRIQIVRASQACWCPRPWDPDRRR-----LTTPAEALVTRRKP	140	
Db	54	AASQWFTYL-TVLARSHRVLAVDLPGRGMSSTPEGPLATGEDAAAFPTAPIEALLSQLAP	112	
Qy	141	-----VILRRATPLQAIAMAMDYDVHLFTDAETGEDAVVYRPGSG	193	
Db	113	GPVAVVGHSGIGGIIVALELALURGVVPERIALVDAMGLGPEMARKAR-----LFFRAGPE-	166	
Qy	184	LRLAROHVPPPGWSCRAPAGPVPL-----IVNSRPTPVLTEAAAVDRAR	230	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:14:10 ; Search time 65.285 Seconds
(without alignments)
1742.576 Million cell updates/sec

Title: US-10-617-038-1

Perfect score: 1419

Sequence: 1 VEPKSRLLVCAPEPSHARE.....YDGNGLITPTGGVADGLA 273

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA.*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1419	100.0	273	15	US-10-617-038-1
2	125.5	8.8	466	14	US-10-156-761-8526
3	108.5	7.6	555	15	US-10-369-493-17959
4	108.5	7.6	1299	16	US-10-437-963-194177
5	108	7.6	937	16	US-10-437-963-127849
6	107.5	7.6	917	15	US-10-282-122A-54089
7	105	7.4	229	15	US-10-425-114-61828
8	103.5	7.3	591	15	US-10-369-493-4549
9	103.5	7.3	823	16	US-10-437-963-128690
10	102.5	7.2	950	16	US-10-437-963-194179
11	102.5	7.2	967	16	US-10-437-963-194213

12	102.5	7.2	19608	15	US-10-084-846A-8	Sequence 8, Appli
13	102	7.2	627	16	US-10-437-963-128801	Sequence 128801,
14	102	7.2	850	16	US-10-437-963-194109	Sequence 194109,
15	102	7.2	1379	16	US-10-437-963-128392	Sequence 128392,
16	101.5	7.2	249	15	US-10-653-595-180	Sequence 180, App
17	101.5	7.2	250	10	US-09-397-945-180	Sequence 180, App
18	101.5	7.2	263	16	US-10-437-963-148780	Sequence 148780,
19	101.5	7.2	283	10	US-09-397-945-333	Sequence 333, App
20	101.5	7.2	283	15	US-10-653-595-333	Sequence 333, App
21	101.5	7.2	516	15	US-10-369-493-7306	Sequence 7306, Ap
22	101.5	7.2	574	14	US-10-156-761-8384	Sequence 8384, Ap
23	101.5	7.2	853	14	US-10-156-761-10186	Sequence 10186, A
24	101	7.1	364	15	US-10-282-122A-61954	Sequence 61954, A
25	101	7.1	613	16	US-10-437-963-187437	Sequence 187437,
26	101	7.1	1024	15	US-10-282-122A-61850	Sequence 61850, A
27	100.5	7.1	5830	17	US-10-211-028-9	Sequence 9, Appli
28	100.5	7.1	19662	15	US-10-084-846A-6	Sequence 6, Appli
29	100	7.0	1164	16	US-10-437-963-193231	Sequence 193231,
30	99.5	7.0	1031	16	US-10-437-963-193435	Sequence 193435,
31	99.5	7.0	1064	16	US-10-437-963-193129	Sequence 193129,
32	99.5	7.0	1209	17	US-10-732-923-8182	Sequence 8182, Ap
33	99	7.0	180	16	US-10-425-115-268653	Sequence 268653,
34	99	7.0	583	16	US-10-437-963-193199	Sequence 193199,
35	99	7.0	939	14	US-10-156-761-7940	Sequence 7940, Ap
36	99	7.0	1974	16	US-10-647-196-24	Sequence 24, Appli
37	98.5	6.9	1267	18	US-10-450-763-53149	Sequence 53149, A
38	98	6.9	1003	16	US-10-437-963-194119	Sequence 194119,
39	97.5	6.9	497	16	US-10-437-963-193273	Sequence 193273,
40	97.5	6.9	1204	16	US-10-437-963-193399	Sequence 193399,
41	97.5	6.9	3668	15	US-10-402-842-4	Sequence 4, Appli
42	97.5	6.9	3668	18	US-10-746-795-4	Sequence 51992, A
43	97	6.8	342	15	US-10-425-114-57689	Sequence 57689, A
44	97	6.8	342	15	US-10-425-114-57689	Sequence 200573,
45	97	6.8	500	16	US-10-437-963-200573	

ALIGNMENTS

RESULT 1

US-10-617-038-1
; Sequence 1, Application US/10617038
; Publication No. US20040057963A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Peter
; APPLICANT: Rosenkrands, Ida
; APPLICANT: Stryhn, Anette
; TITLE OF INVENTION: Therapeutic TB Vaccine
; FILE REFERENCE: SSISAUSA
; CURRENT APPLICATION NUMBER: US/10/617,038
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: DK PA 2002 01098
; PRIOR FILING DATE: 2002-07-13
; PRIOR APPLICATION NUMBER: US 60/401,725
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-617-038-1

Query Match 100.0%; Score 1419; DB 15; Length 273;

Best Local Similarity 100.0%; Pred. No. 1e-129;

Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEPKSRLLVCAPEPSHAREFPDVAVFSGGRANASQAERLARAVGRVLADRGVTGGARVR 60

Db 1 VEPKSRLLVCAPEPSHAREFPDVAVFSGGRANASQAERLARAVGRVLADRGVTGGARVR 60

QY 61 LFWANCADGPTLVQINLVQVGDTPLRQAQATAGIDDLRLPALIRLDROIIVRASQWCPRPWP 120

Db 61 LTMANCADGPTLVQINLQVGDTPFRAQAATAGIDDLRPPALIRLDRQIVRASQWCPRPW 120
QY 121 DRPRRLTTPABALVTRRKPVLRRATPLQATIAAMDAMDYVHLFTDAETGEDAVVYRAG 180
Db 121 DRPRRLTTPABALVTRRKPVLRRATPLQATIAAMDAMDYVHLFTDAETGEDAVVYRAG 180
QY 181 PSGLRLAQHHVFPFGWSCRAPAGPPVPLIVNSRPTPVLTAAAVDRAREHGLPFLFFT 240
Db 181 PSGLRLAQHHVFPFGWSCRAPAGPPVPLIVNSRPTPVLTAAAVDRAREHGLPFLFFT 240
QY 241 DOATRGQLLSRYDGNLGLIPTGADVADGLA 273
Db 241 DOATRGQLLSRYDGNLGLIPTGADVADGLA 273

RESULT 2

US-10-156-761-8526
; Sequence 8526, Application US/10156761
; Publication No. US20030119018A1

; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI

; APPLICANT: IKEDA, HARUO

; APPLICANT: ISHIKAWA, JUN

; APPLICANT: HORIKAWA, HIROSHI

; APPLICANT: SHIBA, TADAYOGHI

; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262

; CURRENT APPLICATION NUMBER: US/10/156,761

; PRIOR FILING DATE: 2002-05-29

; PRIOR FILING DATE: 2001-05-30

; PRIOR FILING DATE: 2001-05-30

; PRIOR FILING DATE: 2001-08-02

; NUMBER OF SEQ ID NOS: 15109

; SEQ ID NO 8526

; LENGTH: 466

; TYPE: PRT

; ORGANISM: Streptomyces avermitilis

US-10-156-761-8526

Query Match 8.8%; Score 125.5; DB 14; Length 466;
Best Local Similarity 24.9%; Pred. No. 0.0024;
Matches 78; Conservative 26; Mismatches 110; Indels 99; Gaps 15;

QY 13 PEPSSHAREFPDVAVFSGGRANASQARLARAVGRVLDRGVGTG-----ARVRLTM 63
Db 34 PQPAQAADLNQLACAVAGTFLAS---RVAEA-----ADKAAAGAAASEDHVVALAAARTAL 85
QY 64 ANCADGPTLVQINLQVGDTPFRAQAATAG-----IDDLRPPALIRLDRQIV 108
Db 86 LGSVHDALMTRVDEIVIGTRGEGATSTPGERHVMVNLAAARSWLCLARAGWQGDHEVV 145
QY 109 RASQWCPRPWDRPRRRRTT-----PAEALVTRRKPV-----VLRRTATLQ 150
Db 146 SASAQVSSMLPDPSSLRLATLTDGFAELASCFGAAL--KMPVRRKGDVWSRAMLLT 203
QY 151 ATAAMD-----MDYDVLFTDAETGEDAVVYRAGPSGLRLARQHVPFGWS 198
Db 204 VPGAADAPTATPACRLLPLGVDVH-----EHATAAQ-----AQHVAVPEADG 247
QY 199 RC-----RAPAGPVP-LIVNS-----RPTPVLTEAAVDRARE-HGLPFLFFTQATCR 246
Db 248 SAPRLVRASVSPKPDVTGVAGWQQLLRPHMSLLAAGEGRSMELHAMPI-----TAE 300
QY 247 GQLLSRYDGNLG 259
Db 301 GDLIWSDEQGRG 313

RESULT 3

US-10-369-493-17959

; Sequence 17959, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 17959

; LENGTH: 555

; TYPE: PRT

; ORGANISM: SPHINGOMONAS

US-10-369-493-17959

Query Match 7.6%; Score 108.5; DB 15; Length 555;

Best Local Similarity 25.9%; Pred. No. 0.14;

Matches 68; Conservative 25; Mismatches 99; Indels 71; Gaps 13;

QY 11 CAPEPSSHAREFPDV-----AVFSGGRANASQARLARAVG-----RV----- 47

Db 5 CSRPAPSSKWTVEIRGIPTRVMKNSPPNAAILARLARGYGRDPTTVYQDERVSYEANY 64

QY 48 -----LADRGVTGARVRLTMANCADGPTLVQINLQVGD--TPLRAQAATA---G 92

Db 65 RATCHLASHMQAQGVAKGDRVALAMRNLPETTFIFAAVSIGAILVPLNAMWTGALEVG 124

QY 93 IDDLRPPALIRLDRQIVRASQWCPRPWDRPR---RRUTTPAEALVTRRKPVLRRATPL 149

Db 125 MKDSGAKLLFVDGERHERLKQ-CYAALPDLRVVVVSRACTPLESEVVALESIV---GPC 179

QY 150 QAIAMD-----AMDYDVLFP-TDAETGEDAVVYRAGPSGLRLARQHVVFPFGWSR 199

Db 180 TGWAALPDIAPEVAIAPDDVTUFTYSGTTGNP-----KGAVGTHRNLMTNLTSGYSV 234

QY 200 CRA-----PAGPPP--VPLIV 212

Db 235 ARAMLRRGEMPPAPPYKVGLLV 257

RESULT 4

US-10-437-963-194177

; Sequence 194177, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 194177

; LENGTH: 1259

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(1259)

```
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_90244C.1.pep
US-10-437-963-194177

Query Match      7.6%; Score 108.5; DB 16; Length 1259;
Best Local Similarity 26.3%; Pred. No. 0.39; 78; Indels 49; Gaps 8;
Matches 52; Conservative 19; Mismatches 19;

QY 83 PLRAQAATAGIDDLRALRLDRQIVRASQWCP--PWPDRPRRLTTPAEALVTRRKP 140
Db 631 PTGVFACTAGLYAAVVGRLRLHRSACXAGDWCFRLHGWPLYRRCWASTSSPTGXLRRR 690
QY 141 VVLRATPQAIAMDYDHLFTDAETGEDAVVYRAGPSGLRLARQHVVFPQWSRC 200
Db 691 LVP---PPARLAFMPPL-LGVHVFTDRXPSPPTGVSASTAGLLCCR-----CWAST 738
QY 201 RAPAG-----BPVPLVNSRPTVLTAAAVDRA---REHGLPFLFFTDQATGRG 247
Db 739 SSTGRVLRRLVFPPLLVGVH-----VFTDLRAFAAANWCFRQHGWPIM----- 783
QY 248 QLLYSRYDGNLGLITPTG 265
Db 784 -----PPLIGLSSPTG 794

RESULT 5
US-10-437-963-127849
; Sequence 127849, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 127849
; LENGTH: 937
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_3025C.1.pep
US-10-437-963-127849

Query Match      7.6%; Score 108; DB 16; Length 937;
Best Local Similarity 28.2%; Pred. No. 0.3;
Matches 55; Conservative 12; Mismatches 58; Indels 70; Gaps 11;

QY 83 PLRAQAATAGIDDLRALRLDRQIVRASQWCP--PWP-----DRP----- 123
Db 422 PTGASACTAGLYAAVVGRLRLHSPAATAADWCFRLHGWPIMLPLGVVYVFTDRPPHPL 481
QY 124 -----RR--RLTTPAEALVTRRKPVLRATPLQAIAMDYDHLFTDAE 169
Db 482 TGVTACTAGLCCRCWASTTPPPACLRRLRVF-----PPARLAFMPPL-LGVYVFT--- 532
QY 170 TGEDAVVYRAGPSGLRLARQHVVFPQWSRC-----APAGPPVPLVNSRPTVLT- 221
Db 533 ---DRLAFAAADWCFRL-----HWPLCRCWASTSSPTGLPSPL-----PTGSAC 576
QY 222 -----EAAAVDRAEH 232
Db 577 TAGLYAAVVGRLRLH 591

; OTHER INFORMATION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54089
; LENGTH: 717
; TYPE: PRT
; ORGANISM: Corynebacterium diptheriae
US-10-282-122A-54089

Query Match      7.6%; Score 107.5; DB 15; Length 717;
Best Local Similarity 22.7%; Pred. No. 0.24;
Matches 62; Conservative 29; Mismatches 107; Indels 75; Gaps 11;

QY 9 VVCAPEPSHAREFPDVAVFSGGRANASQAERLARAVGRVLADRGVTGGARVRLTMANCAD 68
Db 322 LLSAPKQWEQAVDTLRDIGALNAAGQITELGRTLATLPTD-----PRLGASLHLCGG 375
QY 69 G--PTLVQINLV-GDTPLRAQAATAGIDDLRALRLDRQIVRAS-----AQCW 115
Db 376 GAAPTIAALSLDARGDIP-RAIAAMSHQO-----RFTREVARLAFGVDPDSECTGCEAI 427
QY 116 PRPWPDRPRRR-----LTTPAEALVTRRKPVLRATPLQOA 151
Db 428 ARANFKNVARRDGDLSYLMAGGSRFAELGVSSLSGHEWLAVAQASLTHRGSVI-RSAAPISE 486
QY 152 IAAMDAMDYDHLFTDAETGEDAVVYRAGPSGLRLARQHVVFPQWSRCRAPGPVPLI 211
Db 487 AAADVDIIGV-----AERTATLI-----DGSRLGTRIRH-----AGAITLST 523
QY 212 VNSRPTPVLTEAANAVDRAEHGLPFLFFTDQAT 244
```


Db 411 PTGVSACTAGLYAAIVGRLELHRPACXAAADWCFRLHGWFPIMPLLGVVYVFTDRPPRPP 470
QY 129 TPAAE-----LVTRR---KPVLRATPLQAI-----AAMDAM 159
Db 471 TGSACTADLFCRCWASTSSPTGRLVRRRLVFPARLAFMPPLLGVVYVFTDREXASSAAD 530
QY 160 Y-----DVHLFTD-----AETGEDAV-----VYRAGPSGLRLARQH 190
Db 531 WCFRLHGWPLXPLLGVVYVFTDRPPRPPPTGVSACTAGLFCRCWASTSSPTG-RLVRRR 589
QY 191 HVEPPG-----DVHLFTD-----AETGEDAV-----WSECRAPAGPPVPLIYNSR- 215
Db 590 LVFPARLAFMPPLLGVVYVFTDRPPRPPPTGVSACTAGLFCRCWASTSSPTGRLVRRRL 649
QY 216 ---PTPVLTEAAAVDRAREH 232
Db 650 VFPARLAYAAVGRRLH 669

RESULT 10

US-10-437-963-194179
; Sequence 194179, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 194179
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(950)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_90246C.1.pap
US-10-437-963-194179

Query Match 7.2%; Score 102.5; DB 16; Length 950;
Best Local Similarity 27.6%; Pred. No. 1;
Matches 63; Conservative 21; Mismatches 87; Indels 57; Gaps 13;
QY 83 PLRAQAATAGIDDLRLPALIRLDRQIVRASQWCP--PMPDRPRRLTTPAEALVTRRKP 140
Db 492 PTGASACTAGLYATVVGRLRLHRPAAFAADWCFRLHGWPLYRCWASTSSPTGXLRR-- 549
QY 141 VVLRATPLQAIAMDMDYDVHLFTD-----AETGEDAV-----VYRA-----GPSGLR 185
Db 550 --CRLVFPARLAFMPPL-LGVYVFTDRXPPLTGVSACTAGLYAAVXWASTSSPTG-R 605
QY 186 LARQHVVFPFGWSCRAPAGPPVPLIYNSRPTVLTAAAVDRA---REHGLPFP----- 236
Db 606 LRRRLVFPF-----ARLAYXPLLGVI-----VFTDRLAFAAADWCFRLHGWPLPPL 655
QY 237 --LFFTDQ-----ATGRG-----QLLYSRYDGNLGLITPTGDGVA 269
Db 656 GVYVFTDRPPRPPPTGASACTAGLCCRCWASTSSPTGLPSPLPTGV 703

RESULT 11

US-10-437-963-194213
; Sequence 194213, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 194213
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(967)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_90277C.1.pap
US-10-437-963-194213

Query Match 7.2%; Score 102.5; DB 16; Length 967;
Best Local Similarity 28.3%; Pred. No. 1.1;
Matches 51; Conservative 18; Mismatches 74; Indels 37; Gaps 9;
QY 83 PLRAQAATAGIDDLRLPALIRLDRQIVRASQWCP--PMPDRPRRLTTPAEALVTRRKP 140
Db 745 PTGVSACTAGLYAAIVGRRLHRPACXAAADWCFRLHGWPLCRRSWASTSSPTGXLRRR 804
QY 141 VVLRATPLQAIAMDMDYDVHLFTD---AETGEDAVVYRAG-----PSGLR 185
Db 805 LVF-----PPARLAFMPPL-LGVYVFTDRXPPLTGVSACTAGLYAAVXWASTSSPTG-X 858
QY 186 LARQHVVFPFGWSCRAPAGPPVPLIYNSRPTVLTAAAVDRA---REHGLPPLFFTD 241
Db 859 LRRRLVFPF-----ARLAFMPPLGVY-----VFTDRXAAASAAADWCFRLHGWFPVITPTD 908

RESULT 12

US-10-084-846A-8
; Sequence 8, Application US/10084846A
; Publication No. US20040006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MUHLENWEG, AGNES
; APPLICANT: TREFFER, AXEL
; APPLICANT: BECHTHOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: Patentin Ver. 3.2
; SEQ ID NO 8
; LENGTH: 19608
; TYPE: PRT
; ORGANISM: Streptomyces viridochromogenes
; FEATURE:
; OTHER INFORMATION: Protein 3: amino acid sequence encoded by coding strand 1.
; OTHER INFORMATION: Start codon: atc, Start position: nucleotide 3.

US-10-084-846A-8

Query Match 7.2%; Score 102.5; DB 15; Length 19608;
Best Local Similarity 25.1%; Pred. No. 50;
Matches 77; Conservative 15; Mismatches 122; Indels 93; Gaps 15;

QY 11 CAPESSHAREFPDVAVFSGGRANASQAERLA--RAVGRVLADR-----GVTGGARVR 60
DB 13963 CFPGRGRARPRSY-----GPESRPQRRCGCCWRTAGRYGRSRWGPCSAGTTARHGA 14018

QY 61 LTMACAD-----GPTLVQINLVQGTPTLRAQAATAGIDDLRALIRLDRQIVRASQAQ 113
DB 14019 ATRSPAARWRPPSRAGRSRAACSTRAGAHRRSRGPSA---TGRSAAFRSCRR--RRSAR 14073

QY 114 WCPRPWDRPRRLTTPAEALVTRR-KPVULPRATPLQAIAMDAMDYDVHLFTDAETGE 172
DB 14074 WCGSDRPPAGSRNRV--PAR---TRRCSPASGRACRPPGAA-----AGM 14113

QY 173 DAVVYRAGPSGLRLARQHHV-----PPPG-----WSRCRAPA 204
DB 14114 PCAASRPFGP-----ARCSHTAADRIPSAIRPTSSPPAAPPPAPFPFEGGRRRLWRPRRARV 14169

QY 205 GPVPLVINSRTPVLTEAAAVDRAREH-----GLPFLFTDQATGRG-----QLL 250
DB 14170 PRPLRPDRGRVPRPIRRSHTAVRRRREHRCCTGRLSAPTHLPVFRLTGPGAHGSAFFRSSA 14229

QY 251 YSRYDGN 257
DB 14230 FCRYSGN 14236

RESULT 13

US-10-437-963-128801
; Sequence 128801, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 128801
; LENGTH: 627

; ORGANISM: Oryza sativa

; TYPE: PRT

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(627)

; OTHER INFORMATION: unsure at all Xaa locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_3111C.1.pap

US-10-437-963-128801

Query Match 7.2%; Score 102; DB 16; Length 627;

Best Local Similarity 23.9%; Pred. No. 0.69;

Matches 71; Conservative 18; Mismatches 86; Indels 122; Gaps 15;

QY 82 TPLRAQAATAGIDDLRALIRLDRQIVRASAOWCPR--PW-----PDRP- 123

DB 79 SPIRL-ADSAADVGR-----HLHRLAFATAADWCRLYSVWVTPLLGVYVFTDPPRPP 133

QY 124 -----RRRLTTPAEALVTRRKPVVLRRATPLQAIAM 155

DB 134 PTGVSACTAGLLCRCWCWASTSSPTGRLARRLLVFPARLAYAAVVGRLR---LHRPAAS 190

QY 156 DAMDY-----DVHLFTD-----AETGEDAV-----VVRAGBSGLRL 186
DB 191 PAADWCFLHGWPIIMPLLGVYVFTDRPRPPPTGVSACTAGLLCRCWCWAFTS:PTG-RL 249

QY 187 ARQHHVFP-----GMSRCRAPAGPV-----PLIVNSRPTPVLTAAAV 226
DB 250 ARRLVFPFARLAYVAVVGRLLRHRPAASPAADWCFLHGWPIIMPLLGVYVFTXPAAS 309

QY 227 DRA-----REHGLPFL-----FFTQATGRGQLLYSRYDGNL-----GLITPTG 265
DB 310 PAADWCFLHGWPIIMPLLGVYVFTDRPRPPPTGVSAYTASLLCRCWCWASTSSPTG 366

RESULT 14

US-10-437-963-194109
; Sequence 194109, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 194109

; LENGTH: 850

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(850)

; OTHER INFORMATION: unsure at all Xaa locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_90183C.1.pap

US-10-437-963-194109

Query Match 7.2%; Score 102; DB 16; Length 850;

Best Local Similarity 25.6%; Pred. No. 1;

Matches 51; Conservative 14; Mismatches 58; Indels 76; Gaps 10;

QY 83 PLRAQAATAGIDDLRALIRLDRQIVRASAOWCPR--PWP-----DRPR- 125

DB 447 PTGASACTAGPYAAVVGRLRHRPAASAAADWCFLHGWPIIMPLLGVYVFTDRPRPP 506

QY 126 -----RLTTPAEALVTRRKPVVLRRATPLQAIAMDAMDYDVHLFTDA 168

DB 507 TGVSACTAGLLCRCWCWASTSSXPACLRRLRVF-----PPARLAFMPPL-LGVYVFT-- 558

QY 169 ETGEDAVVYRAGPSGLRLAROHVFPFGWSCR-----APAGPPVPLIINSRTPV-- 219

DB 559 ----DRLAFXAADLCFRL-----HGWPLCRCWCWASTSSPTGXP-----RPPTDA 599

QY 220 -----LTEAAAVDRAREH 232

DB 600 SACTAGLAYAAVVGRLRH 618

RESULT 15

US-10-437-963-128392
; Sequence 128392, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 128392
LENGTH: 1379
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(1379)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_3074C.1.pep
US-10-437-963-128392

Query Match 7.2%; Score 102; DB 16; Length 1379;
Best Local Similarity 27.7%; Pred. No. 1.9;
Matches 54; Conservative 13; Mismatches 58; Indels 70; Gaps 11;
QY 83 PLRAQAATAGIDLRPALIRLDRQIVRASAOQWCP--PWP-----DRP----- 123
Db 394 PTCASACTAGLRCRCWASTTTPPACLRRLVF-----PPARLAFMPPL-LGVYVFTDRPPHPL 453
QY 124 -----RR--RLTTPAEALVTRRKPVVLRREATPLQAIAMAMDYDVHLFTDAE 169
Db 454 TGVTTACTAGLRCRCWASTTTPPACLRRLVF-----PPARLAFMPPL-LGVYVFTDRPPHPL 504
QY 170 TGEDAVVYRAGPSGLRLARQHHVFPFGWRCR-----APAGPPVPLVNSRPTPVL- 221
Db 505 ---DRLAFAAADWCFL-----HGWPCLRCRCWASTSPTGLP-----SSLPTGVSA 548
QY 222 -----EAAAVDRAREH 232
Db 549 TAGLYAAVVGRLRLH 563

Search completed: October 15, 2005, 04:14:54
Job time : 66.285 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:59:30 ; Search time 15.2332 Seconds
(without alignments)
1724.342 Million cell updates/sec

Title: US-10-617-038-1
Perfect score: 1419
Sequence: 1 VEPKSRLLVVCAPESPESHARE.....YDGNLGLITPTGGVADGLA 273

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: piri:.*
2: piri2:.*
3: piri3:.*
4: piri4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1416	99.8	273	2 H70849	hypothetical prote
2	120	8.5	835	2 T34770	probable protein p
3	107.5	7.6	203	2 C38179	hypothetical prote
4	106	7.5	705	2 A70669	probable acid-CoA
5	105.5	7.4	559	2 G83024	probable acetolact
6	102.5	7.2	376	2 T55580	adenine deaminase-
7	102	7.2	314	2 AB3610	multidrug resistan
8	101	7.1	452	2 T35762	probable amino aci
9	100.5	7.1	655	2 T35975	probable acyl-pept
10	100.5	7.1	1860	2 T17485	peptide synthetase
11	100	7.0	434	2 G75339	conserved hypothet
12	100	7.0	996	2 C87687	hypothetical prote
13	99	7.0	3643	2 T36410	probable polyketid
14	98	6.9	327	2 JCS319	macrolide-lincosam
15	97	6.8	569	2 E95309	probable calcium b
16	96.5	6.8	242	2 T34951	hypothetical prote
17	96.5	6.8	742	1 TNBBE1	80.7K alpha trans-
18	96	6.8	383	2 C70845	probable nagA prot
19	95.5	6.7	446	2 S32871	regulatory protein
20	95	6.7	336	2 D70794	hypothetical prote
21	95	6.7	509	2 T34871	probable membrane
22	94	6.6	350	2 E70788	hypothetical prote
23	93	6.6	191	2 D75440	probable light-rep
24	93	6.6	268	2 C75416	hypothetical prote
25	92	6.5	1172	2 T36053	probable ABC-type
26	91.5	6.4	290	2 AC3417	DNA polymerase, ba
27	91.5	6.4	404	2 T35421	probable regulator
28	91	6.4	998	2 T35745	probable ATP-bindi
29	90.5	6.4	492	2 JCS169	alkaline nuclease

30	90.5	6.4	1060	2 T31341	ragC protein - Bra
31	90.5	6.4	3670	2 T36249	CDA peptide synthe
32	90	6.3	354	2 B75355	hypothetical prote
33	90	6.3	360	2 D87332	hypothetical prote
34	90	6.3	697	2 T34704	probable transfera
35	89.5	6.3	383	2 T34718	probable integral
36	89.5	6.3	9376	2 T14593	syngomycin synth
37	89	6.3	292	2 T34997	lysr-type transcr
38	89	6.3	708	2 T51026	hypothetical prote
39	89	6.3	3739	2 T17410	polyketide synthas
40	89	6.3	7463	2 T36248	CDA peptide synthe
41	88.5	6.2	308	2 H82212	carboxyphosphonen
42	88.5	6.2	362	2 T36079	hypothetical prote
43	88.5	6.2	477	2 T36111	probable regulator
44	88.5	6.2	948	2 T03225	probable regulator
45	88.5	6.2	958	2 C87504	hypothetical prote

ALIGNMENTS

RESULT 1

H70849
hypothetical protein Rv0079 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: H70849
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: H70849
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-273 <COL>
A;Cross-references: UNIPROT:O53624; GB:AL021428; GB:AL123456; NID:g3261514; PIDN:CAA1626
C;Experimental source: strain H37RV
C;Genetics:
A;Gene: Rv0079
C;Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0079

Query Match	99.8%	Score 1416;	DB 2;	Length 273;
Best Local Similarity	99.6%	Pred. No. 2.3e-111;		
Matches 272;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	VEPKSRLLVVCAPESPESHAREFPDVAVFSGGRANASQAERLARAVGRVLADRGVTGGARVR	60	
Db	1	MEPKSRLLVVCAPESPESHAREFPDVAVFSGGRANASQAERLARAVGRVLADRGVTGGARVR	60	
QY	61	LTWANCADGPTLVQNLQVGTPLRAQAATAGIDDLRALIRLDROI VRASACWCRPPWP	120	
Db	61	LTWANCADGPTLVQNLQVGTPLRAQAATAGIDDLRALIRLDROI VRASACWCRPPWP	120	
QY	121	DRPRRLTTPAALVTRRKPVVLRATPLQAATAAMDYDVHLFTDAETGEDAVVYRAG	180	
Db	121	DRPRRLTTPAALVTRRKPVVLRATPLQAATAAMDYDVHLFTDAETGEDAVVYRAG	180	
QY	181	PSGLRLARQHVVFPFGWSCRAPAGPPVPLI VNSRTPVLTAAAVDRAREHGLPLFFFT	240	
Db	181	PSGLRLARQHVVFPFGWSCRAPAGPPVPLI VNSRTPVLTAAAVDRAREHGLPLFFFT	240	
QY	241	DOATGRGQLLYSRYDGNLGLITPTGGVADGLA	273	
Db	241	DOATGRGQLLYSRYDGNLGLITPTGGVADGLA	273	

RESULT 2

T34770
probable protein pii uridylyltransferase - Streptomyces coelicolor
C;Species: Streptomyces coelicolor

C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C/Accession: T34770
R/Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1998
A/Reference number: Z21557
A/Accession: T34770
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-835 <MUR>
A/Cross-references: UNIPROT:O69873; EMBL:AL023797; PIDN:CAA19377.1; GSPDB:GN00070; SCODEB:GN00070; SCODEB:GN00070; SCODEB:GN00070
A/Experimental source: strain A3(2)
C/Genetics:
A/Gene: glnD; SCODEB:SC2E1.02
C/Superfamily: UTP:GlnB (protein PII) uridylyltransferase

Query Match 8.5%; Score 120; DB 2; Length 835;
Best Local Similarity 25.9%; Pred. No. 0.038;
Matches 82; Conservative 21; Mismatches 109; Indels 104; Gaps 14;

QY 26 VFSGGRANASQAEELARAVGRVLADRGVTGCGARVRL--TMANCADG-----PTLVQIN 76
DB 267 VYZEARVISYAGDVTWREGRVLRSSV----RPLRAMNGNRNGGPKVAERSPLAEV 322
QY 77 LQVQGT-----PLRAQAATAGIDDLRPALIRLDQIVR---ASAQWCPRPW 119
DB 323 EQDGEAVLARTARPERDPALPLRAAAAA-----QAGLPLSRHAVERLAAATARPLTPW 376
QY 120 PDRPRRLTT-----PAEALVTRRKPVVLR-RATPQIAIAAMDAMDYDVH 163
DB 377 PAEAREQLVLLGSGRPTQVWEALEAGLVTRLLPDWERVRCRPNRA-----VH 427
QY 164 LFT-DAETGSDAV-----VVRAGPSGLRLARQHVFPPGWSRCRAPAGPPVLIVNS 214
DB 428 LMTVDRLHIIETAVRAAGFTRRVHR--PDLLIIAALLHDIGKMGWFGDHSVAGETIARDVAA 485
QY 215 R-----PTPVLTEAAAVDRAREHGLPFIFFTDQ-- 242
DB 486 RIGPDGADTAVLATVLRHLLLVETARRDLDDPATVRAVAQAVGTETLELLHALTEAD 545
QY 243 ATGRGQLLYSRYDGNL 258
DB 546 ALATGPAAWSSWRGSL 561

RESULT 3
C38179
C/Species: Bradyrhizobium japonicum
C/Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 09-Jul-2004
C/Accession: C38179
R/Kullik, I.; Fritsche, S.; Knobel, H.; Sanjuan, J.; Hennecke, H.; Fischer, H.M.
J. Bacteriol. 173, 1125-1138, 1991
A/Title: Bradyrhizobium japonicum has two differentially regulated, functional homologs
A/Reference number: A38179; MUID:91123185; PMID:1991712
A/Accession: C38179
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-203 <KUL>
A/Cross-references: UNIPROT:P30334; GB:M59243; NID:g152137; PIDN:AAA26242.1; PID:g152139

Query Match 7.6%; Score 107.5; DB 2; Length 203;
Best Local Similarity 25.8%; Pred. No. 0.084;
Matches 49; Conservative 24; Mismatches 58; Indels 59; Gaps 8;

QY 86 AQAATAGIDDLRPALIRLDQIVRASQACPRPMPDRPRRLTTPAEALVTRRKPVVLR 145
DB 69 APDAYASADQ---ALVMIKELKYKSR-----LKDRSARKAHVASAAL----- 109
QY 146 ATPLOQIAAMDAMDYDVHLFTDAETGEDAVVYRAGPSGLRLARQHHVFPFGWSRCRAPAG 205
DB 110 -----AAMDATSYVLEAPGE-DEDEVT-----GYS----- 135
QY 206 PPVPLIVNSRPTPV--LTEAAAVDRAREHGLPFIFFTDQATGRGQLLYSRVDGNIGLITP 263

C;Accession: G83024
R;Stover, C.K.; Phan, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: G83024
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-559 <STO>
A;Cross-references: UNIPROT:Q9HUI8; GB:AE004910; NID:g9951251; PIDN:AA08036
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA4977
C;Superfamily: acetolactate synthase, large subunit/pyruvate oxidase; thiamin pyrophosph

Query Match 7.4%; Score 105.5; DB 2; Length 559;
Best Local Similarity 25.3%; Pred. No. 0.4;
Matches 72; Conservative 33; Mismatches 113; Indels 67; Gaps 14;
QY 28 SGGANASQAERLARAVGRVLDAG-----VTGGARVRLT--MANCAGDPTLVQINLQ 78
Db 243 AAGEALAAUSERLAAPLFTSVACKGLLPDPAPLNAGASLCVAPGEMIAEADILVAVGT 302
QY 79 VGDT-----PLRAQAATAGID-----DLRPAIIRL---DRQIVRASQWCPRPWDRP 123
Db 303 MADTFWRRLPLSGELIRVDIDPRKFNDFYPSAVALRGDARQTLLEALLVLPQ----- 356
QY 124 RRLTTPAEALVTRREKVVLRATPLQAI--AAMD-----AMDYDVLFTD-----ASTGED 173
Db 357 EARDSAPAAARVARLRAEIRAAHAPLQACHQATLDRIAALPALPADAFVSTDMTQLAYTGN 416
QY 174 AVYVAGPSGLRLARQHHVPPGWSRC--RAPAGPVPLVNSRPTVLT----- 221
Db 417 AFASRAPRSWLH-----PTGYGLGYGLPAGIGAKGAPQPGILVVGOGFLYTAQ 468
QY 222 EAAAVDRAREHGHPFFFTDQATGRQGLLYSRVDGNLGL-ITPTG 265
Db 469 ELATASEELDSPVLWLNNDALQIR-----DDMLGLDIEPVG 507

RESULT 6
C75580
adenine deaminase-related protein - *Deinococcus radiodurans* (strain R1)
C;Species: *Deinococcus radiodurans*
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: C75580
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: C75580
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-376 <WHI>
A;Cross-references: UNIPROT:Q9RYP2; GB:AE001863; GB:AE001825; NID:g6460670; PIDN:AAF1237
A;Experimental source: strain R1
C;Genetics:
A;Gene: DRA0268
A;Map position: 2

Query Match 7.2%; Score 102.5; DB 2; Length 376;
Best Local Similarity 26.5%; Pred. No. 0.45;
Matches 77; Conservative 16; Mismatches 87; Indels 111; Gaps 17;
QY 12 APEPSHAREPPDVAVFSGGRANASQAERLARAVGRVLDAGVTGGARVRLTMANCADGPT 71
Db 77 AGEF---RRVAHPALVSGGR-----ADRAGRA-----GSAAGARYREIL-----GPA 115
QY 72 LVQINLVQVGTGTP-----LRAQAATA---GIDDLR---PALIRLDQIVRASQWCPRPW 119

Db 116 RPR-----GDREPWSSRLRAAAARLATLRGAGNLRGRGSAARRRDAP-----PARW 161
QY 120 PRRPRRL-----TTDAEAL-----VTRRKPVVLRATPLQAI 152
Db 162 RRRPRARLGRHLRSARALAHAPDVSRPDRHRACGAGQRRRPAGRRRPLRARVVELLDV 221
QY 153 AAMDAMDYDVLFTDAE--TGEDAVVYRAGP-----SGLRLARQHHVPPGWSRCRAPAGP- 206
Db 222 GQRPARRHPGHQHSARRASGPGARRQRRGPGAGGSGARTARRH--RPG---RGRRGPR 275
QY 207 -PVPLVNSRPTVLTAAAVDREHGLPFLFTDQATGRQGLLYSRVDG 256
Db 276 AVAPALRGSDRRP-----ASGRGRRCPGPGDG 302

RESULT 7
AB3610
multidrug resistance protein a [imported] - *Brucella melitensis* (strain 16M)
C;Species: *Brucella melitensis*
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AB3610
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, N.
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesee
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*.
A;Reference number: AD3252; PMID:11756688
A;Accession: AB3610
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-314 <KUR>
A;Cross-references: UNIPROT:Q8YBT3; GB:AE008918; PIDN:AAU54045.1; PID:g17984999; GSPDB:G
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BME110803
A;Map position: 11

Query Match 7.2%; Score 102; DB 2; Length 314;
Best Local Similarity 25.8%; Pred. No. 0.4;
Matches 67; Conservative 36; Mismatches 103; Indels 54; Gaps 12;
QY 1 VEPKSRVVVCAPEPSHAREFPDVAVFSG-----RANASQAER-LARA 43
Db 62 VEDTDARIAVAQAQAALQAQAQLANLQVGRPEETAVLEAAVRSARAQADDAQRTLLRT 121
QY 44 VGRVLADRGVTGARV-----RLTMANCADGPTLVQINLVQVDTPLRAQAATAGIDDLRP 98
Db 122 --RDLARRGVATQAQLDDAATQLEVAEAAIGOSTA--NLAVGKLPAPEIKAEANAVKS 177
QY 99 ALIRLDQIVRASQWCPRPWDRPRRLTTTAAELVT---RKPVVLRATPLQAIAM 155
Db 178 AEAQLQ-----TAKW-----QRAQRTTEAPAAGRITDVRNLGDIAGSPAPVLTMLPD 225
QY 156 DAMDYDVLH-----FTDAETGEDAVVYRAGPSGLRLARQHHVFP-PCWSRCRAPAGPVPL 210
Db 226 GAVLKVIYIPEERFSDVAVGSLVSHCDGCGAPGLQARVSVSPDPFT-----BPVIY 278
QY 211 IVNSRPTPV-LTEAAAADRA 229
Db 279 SLETRQKLVLVLEAHPVDPA 298

RESULT 8
T35762
probable amino acid decarboxylase - *Streptomyces coelicolor*
C;Species: *Streptomyces coelicolor*
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C;Accession: T35762
R;Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.
submitted to the EMBL Data Library, August 1999
A;Reference number: Z21588
A;Accession: T35762
A;Status: preliminary; translated from GB/EMBL/DDBJ

RESULT 13

A:Gene: t1rD

C:Superfamily: rRNA (adenine-N6-)-methyltransferase

Query Match 6.9%; Score 98; DB 2; Length 327;

Best Local Similarity 29.8%; Pred. No. 0.91;

Matches 62; Conservative 9; Mismatches 81; Indels 56; Gaps 12;

QY 45 GRVLADRGVTGGARVRLTMAN-----CADGPTLVQINLQVG-DTPLRAQAATAGIDDL 96
DB 88 GPPLADR-LAELEPNLTVVNDFTTWPLPDGFLPFGIGNLPGTGTMLRCLALGPDRC 146
QY 97 RPALIRLRQIVRAS-----AQCPRPWPDPRRRLTTP-----AEALV 135
DB 147 REGLELLQKQYTRKSTGAYGNLFNAQM--EPWYTF--RRLGFPFQEFAPVPGSDTETLL 203
QY 136 TRRKPVLLRRATPLQAIAMDMADYDVHLFTDAETGE--DAVVYRAGSGLELARQHHVF 193
DB 204 VRSRP---RPLAPWSRHAAYQRFVEDVNTSLTIGEAAALDRRAGPGWLGAR----V 256
QY 194 PPG-----WSRCR---APAGPP 207
DB 257 PGLRVKIDITRAEQWERSLPRVAPRRPP 284

RESULT 15

E95309
probable calcium binding transcription regulatory protein [imported] - Sinorhizobium mel
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: E95309
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowe
; Kallman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: E95309
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-569 <KUR>
A:Cross-references: UNIPROT:Q92ZB8; GB:AE006469; PIDN:AAK65039.1; PID:G14523470; GSPDB:C
A:Experimental source: strain 1021, megaplasmid pSymA
R:Galibert, P.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kallman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMA0717
A:Genome: Plasmid

Query Match 6.8%; Score 97; DB 2; Length 569;

Best Local Similarity 22.0%; Pred. No. 2.1;

Matches 68; Conservative 36; Mismatches 103; Indels 102; Gaps 16;

QY 8 LVVCAPEPSHAREFPDVFSGGRANASQAERLARAVGRVLADRGVTGGARVRLTMANCA 67
DB 230 LVVLGP-----ASRLDSSAIHPAGRELMAAARITGTGVAV----AISGPRPRTRSGGFS 281
QY 68 DGPTLVQINLQVGDTP--LRAQAATAGIDDLRALIRLRQIVRASQACWRPWPDPDRPR 125
DB 282 DVQCVLPGWCAQCEAPVWVERKRLYWVDILHPAHRFD-----PVTGKNESCN 330
QY 126 RUTTPAEALVTRKPVVLRATPLQAIAMDMADY-----DVHLFTDAETG-----ED 173
DB 331 VAKLVSAVLPTTRNEGLI-----VASQDGVHEHFDPRGDFNPFPAEPEGLPENRLND 381
QY 174 AVVYRAGPSG-----LRLARQHHVFPFGWSRCRAPAGPVPLIVNSRTPV---LITEAA 224
DB 382 A---KVDPGSLWVGSMRLD-----VSRPTGSLYRLTSAG 413

QY 225 AVDRARE-----HGLPF-----LFFTDQATGRGOLLYS-RVDGNLGLI-----TPT 264
DB 414 EVTRAGSGFTVANGLAWSPDSSTFFVDTVFG---IIYDFDAREGSIANRRVFTVPE 470
QY 265 GDGVADGLA 273
DB 471 AEGRPDGLA 479

Search completed: October 15, 2005, 04:38:18
Job time : 16.2332 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:24:25 ; Search time 61.1503 Seconds
(without alignments)
2286.133 Million cell updates/sec

Title: US-10-617-038-1

Perfect score: 1419

Sequence: 1 VEPKSRLLVCAPEPESHARE.....YDNLGLITPTGGVADGLA 273

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1416	99.8	273	2	O53624 mycobacteri
2	1416	99.8	273	2	Q7U2W3 mycobacteri
3	125.5	8.8	466	2	Q82PD6 streptomyc
4	121	8.5	1863	2	Q7WZ75 nonomuraea
5	120	8.5	835	1	GLND_STRCO
6	113.5	8.0	1567	2	Q63L18 streptomyc
7	111.5	7.9	297	2	Q9RD48 burkholderi
8	110	7.8	618	1	Y4RO_RHISN
9	109	7.7	661	2	Q8RL24 streptomyc
10	108	7.6	937	2	Q8S828 oryza sativ
11	108	7.6	937	2	Q7XYF7 oryza sativ
12	108	7.6	937	2	Q8S828 oryza sativ
13	107.5	7.6	203	1	RP5W_BRAJA
14	107.5	7.6	781	2	Q6NKA3 corynebacte
15	106	7.5	705	2	Q50468 mycobacteri
16	106	7.5	705	2	P96283 mycobacteri
17	106	7.5	705	2	P96283 mycobacteri
18	105.5	7.4	422	2	Q8GH49 mycobacteri
19	105.5	7.4	559	2	Q9HU18 bifidobacte
20	104.5	7.4	4083	2	Q9HU18 pseudomonas
21	104	7.3	334	2	Q63UJ1 nonomuraea
22	104	7.3	1102	2	Q69Q34 burkholderi
23	103.5	7.3	222	2	Q6NIR9 oryza sativ
24	103.5	7.3	551	2	Q9FEE9 corynebacte
25	103.5	7.3	583	2	Q93RW9 streptomyc
26	103	7.3	338	1	YBHG_THIFE
27	102.5	7.2	376	2	Q9RYP2 streptomyc
28	102.5	7.2	751	2	Q62CE5 burkholderi
29	102	7.2	314	2	Q8YBT3 bruceella me
30	102	7.2	421	2	Q8FNF9 corynebacte
31	102	7.2	570	2	Q67TJ5 symbiobacte

32 101.5 7.2 492 2 Q656M4 oryza sativ
33 101.5 7.2 574 2 Q9JGZ7 streptomyc
34 101.5 7.2 853 2 Q82JV5 streptomyc
35 101 7.1 452 2 Q9S2I6 streptomyc
36 101 7.1 513 2 Q69RC5 oryza sativ
37 101 7.1 585 2 Q8PEL4 xanthomonas
38 100.5 7.1 356 2 Q68JD0 escherichia
39 100.5 7.1 575 2 Q93SJ8 myxococcus
40 100.5 7.1 655 2 Q9ZB16 streptomyc
41 100.5 7.1 1121 2 Q98GL0 rhizobium 1
42 100.5 7.1 1188 2 Q73RU2 rhizobium 1
43 100.5 7.1 1860 2 Q5Z821 mycobacteri
44 100 7.0 434 2 Q9RT67 amycolatops
45 100 7.0 996 2 Q9A2M0 deinococcus
Q9A2M0 caulobacter

ALIGNMENTS

RESULT 1

O53624 PRELIMINARY; PRT; 273 AA.
ID O53624
AC O53624: Q7DAH6;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedlocusNames=MT0086, Rv0079;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.F., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekaiia F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrall B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains."
RL J. Bacteriol. 184:5479-5490(2002).
DR EMBL; BX842572; CAA16260.1; -;
DR EMBL; AB000516; AAK44311.1; -;
DR PIR; H70849; H70849.
DR TIGR; MT0086; -;
DR Tuberculist; Rv0079; -;
SQ SEQUENCE 273 AA; 29506 MW; 07F5BDC73A403E2B CRC64;

Query Match 99.8%; Score 1416; DB 2; Length 273;
Best Local Similarity 99.6%; Pred. No. 5.5e-105;
Matches 272; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEPKSRLLVCAPEPESHAREFFPDVAVFSGGRANASQERLARAVGRVLADRGVTGGARVR 60

```
Db 1 MEPKSRLVVCAPESPAREPFDVAVFSGGRANASQAERLARAVGRVLADRGVTGAVR 60
QY 61 LTWACADGPTLVQINLQVGDTPLRQAATAGIDDLRPALIRLDRQIVRASQWCPRPWP 120
Db 61 LTWACADGPTLVQINLQVGDTPLRQAATAGIDDLRPALIRLDRQIVRASQWCPRPWP 120
QY 121 DRPRRLTTPAALVTRRRKPVVLRATPLQAIAAMDAMDYVHLFTDAETGDAVVYRAG 180
Db 121 DRPRRLTTPAALVTRRRKPVVLRATPLQAIAAMDAMDYVHLFTDAETGDAVVYRAG 180
QY 181 PSGLRLARQHVFPPGWSRCRAPGPPVPLVNSRPTPVLTEAAAADRAHREHGLPFLFFT 240
Db 181 PSGLRLARQHVFPPGWSRCRAPGPPVPLVNSRPTPVLTEAAAADRAHREHGLPFLFFT 240
QY 241 DOATRGQLLYSRYDGNLGLIITPTGDGVADGLA 273
Db 241 DOATRGQLLYSRYDGNLGLIITPTGDGVADGLA 273
RESULT 2
Q7U2W3 PRELIMINARY; PRT; 273 AA.
AC Q7U2W3;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Hypothetical protein Mb0082.
GN OrderedLocusNames=Mb0082;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1745;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Dutfoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248334; CAD92944.1; ..
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 273 AA; 29506 MW; 07F5BDC73A403E2B CRC64;

Query Match 99.8%; Score 1416; DB 2; Length 273;
Best Local Similarity 99.6%; Pred. No. 5.5e-105;
Matches 272; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEPKSRLLVVCAPESPAREPFDVAVFSGGRANASQAERLARAVGRVLADRGVTGAVR 60
Db 1 MEPKSRLVVCAPESPAREPFDVAVFSGGRANASQAERLARAVGRVLADRGVTGAVR 60
QY 61 LTWACADGPTLVQINLQVGDTPLRQAATAGIDDLRPALIRLDRQIVRASQWCPRPWP 120
Db 61 LTWACADGPTLVQINLQVGDTPLRQAATAGIDDLRPALIRLDRQIVRASQWCPRPWP 120
QY 121 DRPRRLTTPAALVTRRRKPVVLRATPLQAIAAMDAMDYVHLFTDAETGDAVVYRAG 180
Db 121 DRPRRLTTPAALVTRRRKPVVLRATPLQAIAAMDAMDYVHLFTDAETGDAVVYRAG 180
QY 181 PSGLRLARQHVFPPGWSRCRAPGPPVPLVNSRPTPVLTEAAAADRAHREHGLPFLFFT 240
Db 181 PSGLRLARQHVFPPGWSRCRAPGPPVPLVNSRPTPVLTEAAAADRAHREHGLPFLFFT 240
QY 241 DOATRGQLLYSRYDGNLGLIITPTGDGVADGLA 273
Db 241 DOATRGQLLYSRYDGNLGLIITPTGDGVADGLA 273
RESULT 3
```

```
Q82PD6 PRELIMINARY; PRT; 466 AA.
Q82PD6;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SAV987;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP050325; BAC68697.1; ..
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 466 AA; 48152 MW; 436EEF5AC2E3FD2 CRC64;

Query Match 8.8%; Score 125.5; DB 2; Length 466;
Best Local Similarity 24.9%; Pred. No. 0.097;
Matches 78; Conservative 26; Mismatches 110; Indels 99; Gaps 15;

QY 13 PEPHAREFPDVAVPSGGRANASQAERLARAVGRVLADRGVTGG-----ARVRLTM 63
Db 34 PQAQAADLNQLACAVAGTPLAS---RVAEA-----ADKAAAGAASEDHVALAARTAL 85
QY 64 ANCADGPTLVQINLQVGDTPLRQAATAG-----IDDLRPALIRLDRQIV 108
Db 86 LGSVHDALMTRDEVIGTRGEATASTPGERHMVNLAAASWUCLARAGWQGDHEVV 145
QY 109 RASQWCPRPWDRPRRLTT-----PAEALVTRRRKP-----VLRATPLQ 150
Db 146 SASAQVSSMLPDPRLRLATLLDGFPAELAASCEGAL--KRPVRRWGDVWSRMLLT 203
QY 151 AIAWDA-----MDYDVHLFTDAETGDAVVYRAGPSGLRLARQHVVPPGWS 198
Db 204 VPGAADAPTATAPATGRLLPLGVDVH-----BHATAAQ-----AQHAVPEPADG 247
QY 199 RC-----RAPAGPPV-LIVNS-----RPTPVLTEAAAADRAHREHGLPFLFFTQATGR 246
Db 248 SAPRLVASVSPKPDVTVVAGVWQLLRPHMSLLAAAGEGRSMELHAMPI-----TAE 300
QY 247 QQLLYSRYDGNLGL 259
Db 301 GDLIWSDEQGRPG 313

RESULT 4
Q7WZ75 PRELIMINARY; PRT; 1863 AA.
AC Q7WZ75;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Putative non-ribosomal peptide synthetase, module 7.
```

```

GN Name=dbv16;
OS Nonomurea sp. ATCC 39727.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptosporangineae; Streptosporangiaceae; Nonomurea.
OX NCBI_TaxID=93944;
RN [1]_TaxID=1902;
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 39727.
RX MEDLINE=22721464; PubMed=12837387; DOI=10.1016/S1074-5521(03)00120-0;
RA Sosio M., Stinchi S., Beltrametti F., Lazzarini A., Donadio S.;
RT "The gene cluster for the biosynthesis of the glycopeptide antibiotic
RT A40926 by Nonomurea sp.";
RL Chem. Biol. 10:541-549(2003).
CC -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC family.
DR EMBL; AJ561198; CAD91211.1; -.
DR HSSP; P14687; 1AMU
DR GO; GO:0003824; F: catalytic activity; IEA.
DR GO; GO:0048037; F: cofactor binding; IEA.
DR GO; GO:0016788; F: hydrolase activity, acting on ester bonds; IEA.
DR GO; GO:0009058; P: biosynthesis; IEA.
DR InterPro; IPR010071; AA: adenyl_dom.
DR InterPro; IPR009081; ACP-like.
DR InterPro; IPR000873; AMP-bind.
DR InterPro; IPR001242; Condensatn.
DR InterPro; IPR006163; Phosphanteth_bind.
DR InterPro; IPR000379; Ser. esters.
DR Pfam; PF00501; AMP-binding; 2.
DR Pfam; PF00668; Condensation; 2.
DR Pfam; PF00550; PP-binding; 1.
DR TIGRFAMS; TIGR01733; AA-adenyl-dom; 1.
DR PROSITE; PSS0075; ACP_DOMAIN; 1.
DR PROSITE; PS00455; AMP_BINDING; 1.
DR PROSITE; PS00455; AMP_BINDING; 1.
SQ SEQUENCE 1863 AA; 20054 MW; DIC426E1F4DA07D8 CRC64;

Query Match 8.5%; Score 121; DB 2; Length 1863;
Best Local Similarity 27.2%; Pred. No. 1.1;
Matches 73; Conservative 26; Mismatches 97; Indels 72; Gaps 15;

QY 7 RLVCVAPERSHAREFPDVFVSGGRANASOAE--RLARAVGRVLADRGVTGGARVRLTWA 64
DB 474 RRVATAPDA-----VAITDAGGADLTVAEDQANRLARHLAARGIGRGVRGVWMD 525

QY 65 NCADGETLVQINLQVGDTPLRQAQATAGIDDLRPA-----LIRLDQIVRASQWCPRPWP 120
DB 526 R---SPDLLIAFL-----ASWKAGAAVVPVDVHPAERIEFVLADSGV---SAVLCTR--- 572

QY 121 DRPRRLTTTAAELV-----TRRPVVLRRATPTLQATIAAMDMDYDVLFTDAETG----- 171
DB 573 ---ATREVPADAIVIDAPETRAAIDAGAAATAPQIRLSADDLAY--VMVTSGSTGLPKGV 627

QY 172 ---EDAVV-----VRAGPSGLRLARQHVVFPFG---WSRCRAPGPPVPLIVNSRPT 217
DB 628 GVPHGAVAGUAGDEGRIGPGDAVLVHMATHVFDPSLYAMW-----VPLAMGGR--- 675

QY 218 PVLTEAAAVD-----RAREHGLPFLFFT 240
DB 676 VLTPEGVLDALGMRQAVRGVTFVHLT 703

RESULT 5
GLND_STRCO
ID_GLND_STRCO STANDARD; PRT; 835 AA.
AC O69873;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 25-OCT-2004 (Rel. 45, Last annotation update)
DE (Protein-Pii) uridylyltransferase (BC 2.7.7.59) (Pii uridylyl-
DE transferase) (Uridylvl removing enzyme) (Urase).
GN Name=glnD; OrderedLocusNames=SCO5585; ORFNames=SC2E1.02;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.

```

```

OX NCBI_TaxID=1902;
RN [1]_TaxID=1902;
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21596410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Huang A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)";
RL Nature 417:141-147(2002).
CC -!- FUNCTION: Modifies, by uridylylation or deuridylylation the PII
CC (GlnB) regulatory protein (by similarity).
CC -!- CATALYTIC ACTIVITY: UTP + [protein-Pii] = diphosphate + uridylyl-
CC [protein-Pii].
CC -!- SIMILARITY: Belongs to the glnD family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL939124; CAA19377.1; -.
CC PIR; T34770; T34770.
CC HAMAP; MF 00277; -.
CC InterPro; IPR002912; ACT.
CC InterPro; IPR006674; HD_hydro.
CC InterPro; IPR003607; Met_phos_hydro.
CC InterPro; IPR002934; NTP_transf.
CC InterPro; IPR010043; Urase_glnD.
CC Pfam; PF01842; ACT; 2.
CC Pfam; PF01966; HD; 1.
CC Pfam; PF01909; NTP_transf_2; 1.
CC SMART; SM00471; HDC; 1.
CC TIGRFAMS; TIGR01693; Urase_glnD; 1.
KW Complete proteome; Nucleotidyltransferase; Transferase.
SQ SEQUENCE 835 AA; 89046 MW; 551296D8BE77DE5D CRC64;

Query Match 8.5%; Score 120; DB 1; Length 835;
Best Local Similarity 25.9%; Pred. No. 0.52;
Matches 82; Conservative 21; Mismatches 109; Indels 104; Gaps 14;

QY 26 VFSGGRANASQAERLARAVGRVLADRGVTGGARVRL-TWANCADG-----PTLVQIN 76
DB 267 VTEARVISYAGDVTVREVCVRLRSRV---RPLRAMNGENGCKPVAERSPLAEGVV 332

QY 77 LQVGDT-----PLRAQAATAGIDDLRALIRLDQIVR---ASAQWCPRPW 119
DB 323 EQDGEAVLARTARPERDPAIPLRAAAAAA-----QAGPLSRHVAVRRLAATARPLTPW 376

QY 120 PDRPRRLTT-----PAEALVTRKPVVLR-RATPLQATIAAMDMDYDVH 163
DB 377 PAEAREQLVTLGSGRPTVQVMEALEAGLVRLLLPDWERVRCRPNRA-----VH 427

QY 164 LFT-DAETGEDAV-----VVRAGPSGLRLARQHVVFPFGWSRCRAPGPPVPLIVNS 214
DB 428 LMTVDRHLLETAVRAAGFTRVHR---PDLLIAALLHDIGKPGDHSVAGETIARDVAA 485

QY 215 R-----PTPVLTEAAAVDRAEHGLPFLFTDQ-- 242
DB 486 RIGFDGADTAVLATLVRHLLLVETATRTDLDDPATVRAVAQAVGTETLLELHALTEAD 545

QY 243 ATGRGQLLYSRVDGNL 258

```

```
Db 546 ALATGPAWSSWRGSL 561
RESULT 6
Q63L18 PRELIMINARY; PRT; 1567 AA.
AC Q63L18;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Putative non-ribosomal peptide/polyketide synthase.
GN ORFNames=BPSS1193;
OS Burkholderia pseudomallei K96243.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=272560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K96243;
RX PubMed=15377794;
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdano-Tarraga A.M.,
RA Atkins T., Crossman L.C., Pitt I., Churcher C., Mungall K.,
RA Bentley S.D., Sebaihia M., Thomson N.R., Bacon N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crosset B., Davis P., Deshazer D.,
RA Feltwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitz E., Rutherford K., Sanders M., Simmonds M.,
RA Songvilai S., Stevens K., Tumapa S., Vesaratchavest M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RL Burkholderia pseudomallei."
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245 (2004).
DR EMBL; EX571966; CHA38660.1; --
SQ SEQUENCE 1567 AA; 169000 MW; D285E04CF7E10430 CRC64;

Query Match 8.0%; Score 113.5; DB 2; Length 1567;
Best Local Similarity 29.2%; Pred. No. 3.5;
Matches 66; Conservative 10; Mismatches 63; Indels 87; Gaps 13;

QY 81 DTPLRAQAATAGIDLRPALIRLDR-QIVRASAQCPRPWRPRRRLLTTPAALVTRRK 139
Db 58 DGAALDAQAAFAAWADL-----VDRHQSRLTAFLWERR---DEPRQ----- 94
QY 140 PVVLRRRA-TPLQAI-----AAMDMD--VDVHLFTDAETG-----AVVYRA 179
Db 95 -VVLRRRVPELWDSLDGDDARDADFARLDAHLRADRRRFPAPAPLRLRVAVVRRRA 153
QY 180 GPSGLRLARQHHVFPFGWS-----RCRAPAGPPVPLIVNSRPTPVLTAE 223
Db 154 AARHLMIVSHHLLDGGWSTGVLLGELQALHAARLRRRAALPPAP-----RYA 202
QY 224 AAVDRAREHGLPFLFTTQATGRGOLLVSRDGNGLTTPT---GD 266
Db 203 AFVEHLGRAGRP---CCDAFWRLTL-----AGFAQPTPLVGD 236

RESULT 7
Q9RD48 PRELIMINARY; PRT; 297 AA.
AC Q9RD48;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hypothetical protein SC00635.
GN ORFNames=SCF56.19;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RC STRAIN=A3(2) / M145;

MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RL coelicolor A3(2).";
RL Nature 417:141-147 (2002).
DR EMBL; AL9319106; CAB62764.1; --
GO; GO:0003824; F: catalytic activity; IEA.
DR InterPro; IPR008262; Lipase_AS.
DR InterPro; IPR000379; Ser_estrs.
DR PROSITE; PS00120; LIPASE_SER; UNKNOWN 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 297 AA; 31166 MW; 412FF60F1F9428D1 CRC64;

Query Match 7.9%; Score 111.5; DB 2; Length 297;
Best Local Similarity 25.4%; Pred. No. 0.77;
Matches 71; Conservative 20; Mismatches 98; Indels 91; Gaps 13;

QY 2 EPKESRLVVCAPESPESHAREFPDVAVFSGGRANASQAER-----LARAVGRVLA- 49
Db 50 DPSRDLRLVVRHRRPRT--AAVILHGGRAESRAAPWQALALRMHPFLRALGAATGH 107
QY 50 -----DRGVTCGARVRLTMANCA DGTVLQINLVQDTP-----RAQAA 89
Db 108 DDVLLGQVRYENRCWNGAAEDPLRDARRA-----LAELHGLVGDVPVLLGHSMGRAALR 163
QY 90 TAGIDDLRLPALIRLDRQIVRASAQCPRPWP-----PRRLTTPAE--ALV 135
Db 164 AADADPVRG-----VVALAPWCPGPEVAQLDRALLVLHGSRDVRDPGESAAVY 214
QY 136 TRRRPVVLRATPLQAIAMDMDYDVLHFTDAETGDAVVYRAGPSGLRLAROHVFP 195
Db 215 SRARAAGARAGLL-----LENGDHMLRHH-----ARWHTATA 249
QY 196 GWSRCRAPAGPPVPLI---VNSRTPPVLTAE--AAVDRAR 230
Db 250 AVAHLAPEAAPCELFARALSAAPPVLSRAPPASDAAR 289

RESULT 8
Y4RO_RHISN STANDARD; PRT; 618 AA.
AC P55648;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Hypothetical 69.3 kDa protein Y4RO.
GN ORFNames=Y4RO;
OS Rhizobium sp. (strain NGR234).
OC Plasmid sym. pNGR234a.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Balloch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401 (1997).
CC -!- SIMILARITY: In the N-terminal section; strong to Y4Q.
CC -!- SIMILARITY: In the C-terminal section; some to histidinol-
phosphate aminotransferases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).

CC EMBL; AB000095; AB091840.1; .
 KW Hypochemical protein; Plasmid.
 SQ SEQUENCE 618 AA; 69342 MW; BB01E5847DB3F53C CRC64;

Query Match 7.8%; Score 110; DB 1; Length 618;
 Best Local Similarity 22.1%; Pred. No. 2.3;
 Matches 73; Conservative 37; Mismatches 106; Indels 114; Gaps 15;

QY 4 KRSRLVV-CAPEPSHAR-----EFPDVAVFG-----GRANSAQERLAPAVGRVLA 49
 DB 73 KRMELVIRIMPTGSGRLDLATITBIAEPHGQVLVFTENAVIDPSLIQRLRS-----S 127
 QY 50 DRGVT-----GGARVLTMCACDPTLVQINLOVGDTPPLRAQATAGIDD 95
 DB 128 DRNITVVSRRRCRLRLGIDIGRLTA-----MLPGDSPRESASA-----D 170
 QY 96 LRP-ALIRLDQIVRASACQCPWPDPRPERRLTTPAEALVTRKPVVLRRTATPLQAI 154
 DB 171 VSPVGKFKPDALIRAIRPRFND-----DLEFFETALGLQRHQIYLMYADPQHVRRV 226
 QY 155 MDAMDYDVHLFTDAETGEDAV---YRAG-----PSGLRLARQHHVFPFGWSRCRAPAGPP 207
 DB 227 NDATDLEAANFSSSGDQFSILERLQAGNWRYPASEHILLCNHFFP----- 274
 QY 208 VPLVNSRPTPVTEAAVDRARE--HGLPFLFTDQ-----ATGRG 247
 DB 275 -----ASVVDRLERLQDLVLYQPSDQDIIAKLSEMTDLPARNLAVGNG 319
 QY 248 -----OLLYSRVGNGLIPTGCGVADGL 272
 DB 320 VGEILKALVGLDPRIVPTPTSAQYIDAV 349

RESULT 9
 Q8LS06 PRELIMINARY; PRT; 1001 AA.
 ID Q8LS06
 AC Q8LS06
 DT 01-OCT-2002 (T-EMBLrel. 22, Created)
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
 DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
 DE Putative gag-pol polyprotein.
 GN ORFNames=OSUNB0091B12.16;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
 RA Sasaki C., Henry D., Oates R., Simmons J.;
 RA Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA The Rice Chromosome 10 Sequencing Consortium;
 RA "In-depth view of structure, activity, and evolution of rice
 RT chromosome 10."
 RN Science 300:1566-1569(2003).
 RP [3]
 RP SEQUENCE FROM N.A.
 RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC093178; AAM22020.1; .
 DR EMBL; AE017070; AAP2696.1; .
 DR Gramene; Q8LS06; .
 DR InterPro; IPR005213; HGWP.
 DR InterPro; IPR008916; Retrov_capsid_C.

DR Pfam; PF03578; HGWP; 9.
 KW Polyprotein.
 SQ SEQUENCE 1001 AA; 112260 MW; 80D2958C24BEFD54 CRC64;
 Query Match 7.8%; Score 110; DB 2; Length 1001;
 Best Local Similarity 28.7%; Pred. No. 4;
 Matches 56; Conservative 11; Mismatches 58; Indels 70; Gaps 11;
 QY 83 PLRAQAATAGIDDLRALIRLDRQIVRASACQCPR--PMP-----DRP----- 123
 DB 602 PTGASACTAGLAAVAVVGRILHRPAASTAADCFLHGWPIMLPLLGVVVFTDRPPHPPL 661
 QY 124 -----RR--RLTTPAEALVTRKPVVLRRTATPLQAIADAMDYDVHLFTDAE 169
 DB 662 TGVTACTAGLLCCRCRCWASTTTPPACLHRRRLVF-----PPARLAFMPPL-LGVVFT--- 712
 QY 170 TGEDAVVYVAGPSGLRLARQHHVFPFGWSRCR-----APAGPPVPLVNSRPTVLT- 221
 DB 713 ---DRLAFAAADWCFL-----HGWPLCRCWASTSPTGLPSPL-----PTGVSA 756
 QY 222 ----EAAAVDRAREH 232
 DB 757 TAGLYAAVVGRLRLH 771
 RESULT 10
 Q8RZL4 PRELIMINARY; PRT; 661 AA.
 ID Q8RZL4
 AC Q8RZL4
 DT 01-JUN-2002 (T-EMBLrel. 21, Created)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
 DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
 DE B1168H06.23 protein.
 GN Name=B1168H06.23;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=12447439; DOI=10.1038/nature01184;
 RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
 RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
 RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
 RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
 RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
 RA Ikono M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
 RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
 RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
 RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
 RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K., Shibata M.,
 RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
 RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
 RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
 RA Yano M., Jiang J., Gojobori T.;
 RT "The genome sequence and structure of rice chromosome 1";
 RL Nature 420:312-316(2002).
 DR EMBL; AP03563; BAB89182.1; .
 DR Gramene; Q8RZL4; .
 DR InterPro; IPR005213; HGWP.
 DR Pfam; PF03578; HGWP; 7.
 SQ SEQUENCE 661 AA; 74940 MW; 2EE21E51308B6C91 CRC64;
 Query Match 7.7%; Score 109; DB 2; Length 661;
 Best Local Similarity 28.5%; Pred. No. 3;
 Matches 63; Conservative 22; Mismatches 74; Indels 62; Gaps 13;
 QY 101 IRLDQIVRASACQCPR--PWPDRPRRLTTTAAEALVTRKPVVLRRTATPLQAIADAM 158
 DB 330 LRLHRPAAFAAADWCFLHGWPLCRCWASTSPTGLRRRLRV---PPARLAFMPPL 385
 QY 159 DVDVHLFTD-----AETGEDA-----VVYR-----AGPSGLRLARQHHVFPFGWSRCRAP 203

Db 366 -LGVIYVFTDRPPPTGASACTAGLLCYRCWASTSSPTG-RLRRRLVFP-ARLA 439
 QY 204 AGPP---VPLIVNSRPTPVLT-----EAAVDR-----REHGLPF 236
 Db 440 FMPPLLGVIYVFTDRPPPTGASACTAGLYAAVVGRLRHRAAASSAADWCRHLHGWI 499
 QY 237 L-----PFTDOATCRGOLLYSRVGNLGLTPTGQGVAD 270
 Db 500 MPPLLGVIYVFTDRPP-HPPLTATTDEHRLSPSTGYAVAD 539

RESULT 11

Q8S828 PRELIMINARY; PRT; 937 AA.
 AC Q8S828;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Hypothetical protein OSJNBa0053D03.29.
 GN Name=OSJNBa0053D03.29;
 OS Oryza sativa (rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1];
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nipponbare;
 RA McCombie W.R., Cordum H., Minx P., de la Bastide M., Spiegel L.,
 RA Nascimben L., Zucavern T., Balija V., Bell M., Baker J., Santos L.,
 RA Miller B., Katzenberger P., Muller S., Preston R., Kirchoff K.,
 RA Kuit K., King L., Yang C., O'Shughnessy A., Palmer L., Dedhia N.,
 RA Wilton R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC021892; AA00965.1; -;
 DR Gramene; Q8S828; -;
 DR InterPro; IPR005213; HGWP.
 DR Pfam; PF03578; HGWP; 11.
 KW Hypothetical protein.
 SQ SEQUENCE 937 AA; 103904 MW; 4C0C586A18B96FDD CRC64;

Query Match 7.6%; Score 108; DB 2; Length 937;
 Best Local Similarity 28.2%; Pred. No. 5.3;
 Matches 55; Conservative 12; Mismatches 58; Indels 70; Gaps 11;

QY 83 PLRAQAATAGIDDLRALIRLDQIVRASAOQCPR--PWP-----DRP----- 123
 Db 422 PTGASACTAGLYAAVVGRLRHRAAASSAADWCFRLHGWPIMPLLGVIYVFTDRPPHPL 481
 QY 124 -----RR--RLTTPAEALVTRKPVVLRATPLQAIAMDAMDYDVHLFTDAE 169
 Db 482 TGVACTAGLLCCRCWASTTTPPACLHRRRLVF-----PPARLAFMPPL-LGVYVFT--- 532
 QY 170 TGEDAVVYVAGPSGLRLARQHVPFGWSRCR-----APAGPVPLIVNSRPTPVLT- 221
 Db 533 ---DLAFAAADWCFL-----HGWPLCRCWASTSSPTGLSP-PTGVSAC 576
 QY 222 ----EAAAVDRAREH 232
 Db 577 TAGLYAAVVGRLRLH 591

RESULT 12

Q7XFY7 PRELIMINARY; PRT; 937 AA.
 AC Q7XFY7;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=OSJNBa0053D03.29;
 OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 RN [1];
 RP SEQUENCE FROM N.A.
 RA "In-depth view of the structure, activity, and evolution of rice
 chromosome 10.";
 RL Science 300:1566-1569(2003).
 RN [2];
 RP SEQUENCE FROM N.A.
 RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017074; AAP52897.1; -;
 DR Gramene; Q7XFY7; -;
 DR InterPro; IPR005213; HGWP.
 DR Pfam; PF03578; HGWP; 11.
 KW Hypothetical protein.
 SQ SEQUENCE 937 AA; 103904 MW; 4C0C586A18B96FDD CRC64;
 Query Match 7.6%; Score 108; DB 2; Length 937;
 Best Local Similarity 28.2%; Pred. No. 5.3;
 Matches 55; Conservative 12; Mismatches 58; Indels 70; Gaps 11;

QY 83 PLRAQAATAGIDDLRALIRLDQIVRASAOQCPR--PWP-----DRP----- 123
 Db 422 PTGASACTAGLYAAVVGRLRHRAAASSAADWCFRLHGWPIMPLLGVIYVFTDRPPHPL 481
 QY 124 -----RR--RLTTPAEALVTRKPVVLRATPLQAIAMDAMDYDVHLFTDAE 169
 Db 482 TGVACTAGLLCCRCWASTTTPPACLHRRRLVF-----PPARLAFMPPL-LGVYVFT--- 532
 QY 170 TGEDAVVYVAGPSGLRLARQHVPFGWSRCR-----APAGPVPLIVNSRPTPVLT- 221
 Db 533 ---DLAFAAADWCFL-----HGWPLCRCWASTSSPTGLSP-PTGVSAC 576
 QY 222 ----EAAAVDRAREH 232
 Db 577 TAGLYAAVVGRLRLH 591

RESULT 13
 RPSM_BRAJA STANDARD; PRT; 203 AA.
 AC P30334;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Probable sigma(54) modulation protein (ORF203).
 GN OrderedLocusName=blr0724;
 OS Bradyrhizobium japonicum.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae; Bradyrhizobium.
 OX NCBI_TaxID=375;
 RN [1];
 RP SEQUENCE FROM N.A.
 RC STRAIN=USDA 1109pc4;
 RX MEDLINE=91123185; PubMed=1991712;
 RA Kullik I., Fritsche S., Knobel H., Sanjuan J., Hennecke H.,
 RA Fischer H.-M.;
 RT "Bradyrhizobium japonicum has two differentially regulated, functional
 homologs of the sigma 54 gene (rpoN).";
 RL J. Bacteriol. 173:1125-1138(1991).
 RN [2];
 RP SEQUENCE FROM N.A.
 RC STRAIN=USDA 110;
 RX MEDLINE=22484998; PubMed=12597275;
 RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
 RA Sasamoto S., Watanabe A., Idenawa K., Iiiguchi M., Kawashima K.,
 RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
 RA Tabata S.;
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium


```
Query Match      7.5%; Score 106; DB 2; Length 705;
Best Local Similarity 24.6%; Pred. No. 5.6;
Matches 65; Conservative 31; Mismatches 104; Indels 64; Gaps 10;

Qy 12 APEPSHAREFPDVAVFSGGRANASQAERLARAVGRVLADRGVTGGARVRLTMANCADGPT 71
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 11 AEQASEAGWYDRPAFYAADVVTGQIHGGAARLGEVLRNRGLSSGDRVLCL---PDSPD 67
   ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 72 LVQINLVQGDTPLRQAATAGIDDLRPAIRLDROIIVRASAOQWCPRPWDRPRRLTTTPA 131
   ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 68 LVQLLLACLARGVMAFLA-----NPELHRDDHALAARNT-----PALVVT 109
   ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 132 EALVTRRKFPVLRRAATPLQAIAM-----DAMDY-----DVHLFTD 167
   ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 110 DALRDRFQPSRVAEAAELMSEARVAPGGYEPMGGDALAYATYTGTPPKAAIHRHAD 169
   ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 168 AETGEDAVVYRA-----GPSGLRLARQHVFPPG---MSRCRAPAGPPVPLIVNSRPTPV 219
   ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 170 PLTFVDAMCRKALRLTPEDTGLCSARMYPAYGLGNSVM---FPLATGGSAVINS--APV 223
   ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 220 LTEAAVDRAR-----EHGLPELF 238
   ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 224 TPEAAAILSAFPGPSVLYGVNPF 247
   ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
```

Search completed: October 15, 2005, 04:35:47
Job time : 63.1503 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:20:35 ; Search time 40.4081 Seconds
(without alignments)
1454.845 Million cell updates/sec

Title: US-10-617-038-2

Perfect score: 755

Sequence: 1 MDPGSRASQSAREVVELD.....VAIEPIVTGIRIVADSRTP 152

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	755	100.0	152	8	ADI37281 M. tuberc
2	131	17.4	157	4	Aau50385 Propionib
3	131	17.4	157	6	Abm46904 Propionib
4	100.5	13.3	97	5	Abp01079 Human ORF
5	85.5	11.3	192	8	AdS30737 Bacterial
6	85.5	11.3	654	4	Aau34768 E. coli c
7	85.5	11.3	654	6	Aau28792 Protein e
8	84.5	11.2	11096	4	Aae10129 Streptomy
9	83	11.0	304	8	AdN25437 Bacterial
10	82.5	10.9	110	8	ADI37317 M. tuberc
11	82.5	10.9	408	6	ABU33917 Protein e
12	82.5	10.9	1386	7	ABO79841 Pseudomon
13	82	10.9	454	6	ABU47143 Protein e
14	82	10.9	454	6	ABU47772 Protein e
15	82	10.9	566	5	Aau76437 Novel des
16	81.5	10.8	451	6	ABU32163 Protein e
17	81.5	10.8	505	7	ABO67533 Klebsiell
18	81.5	10.8	577	6	ABU22594 Protein e
19	81	10.7	445	6	ABU45054 Protein e
20	80.5	10.7	204	7	ABO80479 Pseudomon
21	80.5	10.7	329	7	ADB65555 Human pro
22	80.5	10.7	3931	6	ABU07377 Human pro
23	80.5	10.7	3931	7	ADG39786 Human nov
24	80	10.6	518	8	ADN25663 Bacterial
25	80	10.6	1180	8	ADN25663 Bacterial

26	79.5	10.5	1490	7	ABO79696 Pseudomon
27	79	10.5	305	7	ADM26140 Hyperther
28	79	10.5	448	6	ABU34774 Protein e
29	79	10.5	448	6	ABU36933 Protein e
30	78.5	10.4	133	5	ABP00814 Human ORF
31	78.5	10.4	340	7	ADM25413 Hyperther
32	78.5	10.4	919	8	ADN22714 Bacterial
33	78.5	10.4	994	5	ABP62769 Protein f
34	78.5	10.4	984	7	ADJ72180 Streptomy
35	78	10.3	288	8	ADI39298 S. hygro
36	78	10.3	386	8	ADR28226 Streptomy
37	78	10.3	813	8	ADN20951 Bacterial
38	77.5	10.3	160	6	ABU35412 Protein e
39	77.5	10.3	163	8	ADL04571 M. catarr
40	77.5	10.3	728	5	ABP65783 Bifidobac
41	77	10.2	288	8	ADI39266 S. hygro
42	77	10.2	366	8	ADS24730 Bacterial
43	77	10.2	1144	7	ABO77842 Pseudomon
44	77	10.2	1432	8	ADQ62813 Pseudomon
45	77	10.2	1612	2	AAR66457 Chimeric

ALIGNMENTS

RESULT 1
ADI37281
ID ADI37281 standard; protein; 152 AA.
XX
AC ADI37281;
XX
DT 22-APR-2004 (first entry)
XX
DE M. tuberculosis low oxygen induced antigen Rv0080 SEQ ID NO:2.
XX
KM mycobacterial infection; vaccine; tuberculosis;
KW Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;
KW low oxygen induced antigen.
XX
OS Mycobacterium tuberculosis.
XX
PN WO2004006952-A2.
XX
PD 22-JAN-2004.
XX
PF 08-JUL-2003; 2003WO-DK000477.
XX
PR 13-JUL-2002; 2002BK-00001098.
XX
PA (STAT-) STATENS SERUM INST.
XX
PI Andersen P, Rosenkrands I, Stryhn A;
XX
DR WPI; 2004-1227778/12.
DR N-PSDB; ADI373326.
XX
PT Use of one or more polypeptides or their fragments, which are expressed during the latent stage of the mycobacterial infection, and/or nucleic acids encoding the polypeptides, for a therapeutic vaccine against tuberculosis.
XX
PS Claim 3; SEQ ID NO 2; 76pp; English.
XX
CC The present invention describes polypeptides or their fragments, which are expressed during the latent stage of a mycobacterial infection, and/or nucleic acids encoding the polypeptides, which are useful for creating a therapeutic vaccine against tuberculosis. Also described: (1) a therapeutic vaccine against tuberculosis comprising one or more polypeptides; (2) a method for treating an animal, including a human being, with tuberculosis caused by virulent mycobacteria, e.g. by Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for immunising an animal, including a human being, against tuberculosis caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis

CC caused by virulent mycobacteria in an animal, including a human being;
 CC (5) a method for diagnosing previous or ongoing infection with a virulent
 CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
 CC infection in a subject. The polypeptides have antibacterial activities,
 CC and can be used in vaccines and in gene therapy. The polypeptides are
 CC useful for the manufacture of a therapeutic vaccine for treating an
 CC individual who is infected by a virulent mycobacterium, e.g. M.
 CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
 CC The present sequence represents a low oxygen induced antigen, which is
 CC used in the exemplification of the present invention.

XX
 SQ Sequence 152 AA;

Query Match 100.0%; Score 755; DB 8; Length 152;
 Best Local Similarity 100.0%; Pred. No. 4.1e-79;
 Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPGRRASPOQAREVVELDRDEAMRLLASVDHGRVVFTRAAALPAIRPNVHLLVVDGRVIG 60

DB 1 MSPGRRASPOQAREVVELDRDEAMRLLASVDHGRVVFTRAAALPAIRPNVHLLVVDGRVIG 60

QY 61 RTRLTAKSVAVRSSADAGVVVAYEADDLPDRRTGWSVVVTGLATEVSDPEQVARYQRL 120

DB 61 RTRLTAKSVAVRSSADAGVVVAYEADDLPDRRTGWSVVVTGLATEVSDPEQVARYQRL 120

QY 121 LHPWNMAMDTVVAIEPEIVTGTGIRIVADSRTP 152

DB 121 LHPWNMAMDTVVAIEPEIVTGTGIRIVADSRTP 152

RESULT 2

AAU50385

ID AAU50385 standard; protein; 157 AA.

AC AAU50385;

XX 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #11281.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-0199047P.

XX 02-JUN-2000; 2000US-0208841P.

XX 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI: 2001-616774/71.

XX N-PSDB; AAS59348.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.

PS Claim 6; SEQ ID NO 11580; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in

CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 157 AA;

Query Match 17.4%; Score 131; DB 4; Length 157;
 Best Local Similarity 29.8%; Pred. No. 1e-06;
 Matches 42; Conservative 27; Mismatches 60; Indels 12; Gaps 6;

QY 6 RASPOSARE--VVELDRDEAMRLLASVDHGRVVFTRAAALPAIRPNVHLLVVDGR-VIGRT 62

DB 13 RKNRPMEINEENPITEVPEDSCWYLDTAEVGRLATSVNDQPEVFPVNV-VDQSQIVFRT 71

QY 63 RLTAKSVAVRSSADAGVVVAYEADDLPDRRTGWSVVVTGLATEVSDPEQVARYQRL-L 121

DB 72 AAGSKLEDIVTNNR-----VAFEADGW--TEEAGSVVLRGKAEIITDDAELALCDKMPL 124

QY 122 HPWNMAMDTVVAIEPEIVTG 142

DB 125 LPWPTVKRNVVRVEADQITG 145

RESULT 3

ABM46904

ID ABM46904 standard; protein; 157 AA.

AC ABM46904;

XX 20-OCT-2003 (first entry)

XX Propionibacterium acnes immunogenic polypeptide #11580.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
 KW immunostimulant; immune response; vaccine; immunogenic.

XX Propionibacterium acnes.

XX WO2003033515-A1.

XX 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

XX 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

XX Zhang Y, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

XX Barth B, Vailieve-Douglass J;

XX WPI: 2003-381789/36.

XX N-PSDB; ACF64477.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.

XX PS Claim 6; SEQ ID NO 11580; 1481bp; English.

XX CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)

CC encoding a Propionibacterium acnes protein. The invention also relates to

CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to

CC immunogenic fragments of P. acnes polypeptides. The invention

CC additionally encompasses expression vectors and host cells comprising a

CC polynucleotide of the invention; antibodies against polypeptides of the

CC invention; fusion proteins comprising a polypeptide of the invention; a

CC method for stimulating an immune response specific for a P. acnes

CC polypeptide and an isolated T cell population comprising T cells prepared

CC via this method; a vaccine composition (comprising P. acnes polypeptides,

CC polynucleotides, antibodies, fusion proteins, T cell populations, or

CC antigen-presenting cells that express the polypeptide); a method and kit

CC for detecting or determining the presence or absence of P. acnes in a

CC patient; and a method for inhibiting the development of P. acnes in a

CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion

CC proteins, T cell populations or antigen-presenting cells that express the

CC polypeptides are useful for diagnosing, preventing or treating acne

CC vulgaris, or for stimulating an immune response specific for a P. acnes

CC protein. The polynucleotides can also be used as probes or primers for

CC nucleic acid hybridisation. The vaccine composition is useful for the

CC stimulation of an immune response against P. acnes, or for treating acne,

CC and the kit is useful for performing a diagnostic assay. The present

CC sequence represents a specifically claimed P. acnes polypeptide which is

CC thought to contain an immunogenic region. Note: The sequence data for

CC this patent did not form part of the printed specification, but was

CC obtained in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX CC Sequence 157 AA;

XX SQ

Query Match 17.4%; Score 131; DB 6; Length 157;

Best Local Similarity 29.8%; Pred. No. 1e-06;

Matches 42; Conservative 27; Mismatches 60; Indels 12; Gaps 6;

QY 6 RRASPOSARE--VVELDRDEAMLLASVDHGRVVFTRAAALPAIRPNYHLVDGR-VIGRT 62

DB 13 RKNRPINEENPITEVPEDSCWGLDTAEVGRLATSVNDQPEFPVNY-VVDQSIVPRT 71

QY 63 RLTAQSVAVRSSADAGVAVVAYEADDLPDRRTGWSVVVTGLATEVSDPEQVARYQL-L 121

DB 72 AAGSKLEDIVTNRR-----VAFEDAGW--TEEAGWSVVLGRKAEIITDDAELALCDKMP 124

QY 122 HPWVNMAMDTVVAIEPIVTG 142

DB 125 LPWVPTVKRNVVRVEADQITG 145

RESULT 4

ABP01079

XX ID ABP01079 standard; protein; 97 AA.

XX AC ABP01079;

XX XX

XX 24-JUN-2002 (first entry)

XX DE Human ORFX protein sequence SEQ ID NO:2140.

XX XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;

XX hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;

XX degenerative disorder; osteoarthritis; neurodegenerative disorder;

XX cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;

XX hypertension; hypothyroidism; cholesterol ester storage disease;

XX immune deficiency; immune disorder; infectious disease;

XX autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;

XX myasthenia gravis.

XX OS Homo sapiens.

XX XX WO200192523-A2.

XX PN

XX XX

106-DEC-2001.

29-MAY-2001; 2001WO-US010836.

30-MAY-2000; 2000US-0206132P.

29-AUG-2000; 2000US-0228716P.

(CURA-) CURAGEN CORP.

Shimkets RA, Leach MD;

WPI; 2002-106308/14.

N-PSDB; ABN16831.

Novel human polypeptides and polynucleotides useful for diagnosing,

PT preventing and treating cardiovascular disease, neurodegenerative,

PT hyperproliferative disorders and autoimmune disorders.

PS Disclosure; SEQ ID NO 2140; 1037pp; English.

XX CC The present invention describes substantially purified human proteins

CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1

CC in the specification). ABN15762 to ABN27252 encode the human ORFX

CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for

CC treating or preventing a pathology associated with an ORFX-associated

CC disorder in humans, and in the manufacture of a medicament for treating a

CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide

CC sequences can be used in gene therapy. ORFX sequences can be used in the

CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,

CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,

CC osteoarthritis, neurodegenerative disorders, disorders related to organ

CC transplantation, cardiovascular diseases, diabetes mellitus, systemic

CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester

CC storage disease, various immune deficiencies and disorders, rheumatoid

CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid

CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host

CC disease and autoimmune inflammatory eye disease. ORFX proteins are also

CC useful for treating burns, incisions, ulcers, for treating osteoporosis,

CC bone degenerative disorders, or periodontal disease, and for gut

CC protection or regeneration and treatment of lung or liver fibrosis,

CC reperfusion injury in various tissues and conditions resulting from

CC systemic cytokine damage. N.B. The sequence data for this patent did not

CC form part of the printed specification, but was obtained in electronic

CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX CC Sequence 97 AA;

XX SQ

Query Match 13.3%; Score 100.5; DB 5; Length 97;

Best Local Similarity 33.7%; Pred. No. 0.0018;

Matches 31; Conservative 16; Mismatches 36; Indels 9; Gaps 4;

QY 53 VVDGRVIG-RTRLTAQSVAVRSSADAGVAVVAYEADDLPDRRTGWSVVVTGLATEVSDP 111

DB 1 VVDGSIAPRTAAGSKLEDIVTNRR-----VAFEDAGW--TEEAGWSVVLGRKAEIITDD 53

QY 112 EQVARYQL-LHPWVNMAMDTVVAIEPIVTG 142

DB 54 AELALCDKMPLLPWVPTVKRNVVRVEADQITG 85

RESULT 5

ADS30737

ID ADS30737 standard; protein; 192 AA.

XX AC ADS30737;

XX XX

XX 02-DEC-2004 (first entry)

XX DT Bacterial polypeptide #19770.

XX DE

XX XX Recombinant DNA construct; transformed plant; improved plant property;

XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;

XX pathogen tolerance; pest tolerance; plant disease resistance;

KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX Bacteria.
XX US2003233675-A1.
XX 18-DEC-2003.
XX 20-FEB-2003; 2003US-00369493.
XX 21-FEB-2002; 2002US-0360039P.
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX Claim 1; SEQ ID NO 19770; 122pp; English.
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 192 AA;
SQ
Query Match 11.3%; Score 85.5; DB 8; Length 192;
Best Local Similarity 31.4%; Pred. No. 0.25;
Matches 33; Conservative 15; Mismatches 24; Indels 33; Gaps 6;
QY 65 TAKVSVVRSSADAGV-----VVAYEADDLDPRR-----TCMSVVVTGL 104
DB 3 TSKPEVVITGASGIGRATVRAFAKRGAYIGLVARSRDGLEAAQVEEAAGAKALV--L 60
QY 105 ATEVSDPEQV-----ARYQRLHP-----WNNMAMDTVVA-----IEPE 138
DB 61 PTDVSDPEQVEAAAAAVERKEFGPIDIWNVDAMASILSPFTEIKPE 105
RESULT 6
AAU34768

ID AAU34768 standard; protein; 654 AA.
XX
AC AAU34768;
XX
DT 14-FEB-2002 (first entry)
XX
DE E. coli cellular proliferation protein #349.
XX
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW antibacterial; drug design.
XX
XX Escherichia coli.
XX OS
XX WO200170955-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US009180.
XX
XX 21-MAR-2000; 2000US-0191078P.
XX 23-MAY-2000; 2000US-0206848P.
XX 26-MAY-2000; 2000US-0207727P.
XX 23-OCT-2000; 2000US-0242578P.
XX 27-NOV-2000; 2000US-0253623P.
XX 22-DEC-2000; 2000US-0257931P.
XX 16-FEB-2001; 2001US-0269308P.
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
XX N-FSDB; AAS52627.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Example 3; SEQ ID NO 10361; 511pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 654 AA;
SQ
Query Match 11.3%; Score 85.5; DB 4; Length 654;
Best Local Similarity 27.1%; Pred. No. 1.4;
Matches 36; Conservative 20; Mismatches 64; Indels 13; Gaps 4;
QY 20 DRDEAMRLLASVDHGR---VVFTRAALPAIRPVNHLVVDGRVIGTRITAKV-SVAVRSS 75
DB 208 DLAEITNOLISESHGKSQMPALLSAKIVADKRTNSLIISGPEKARQRTISLLKSLDVEES 267
QY 76 ADAGVVVAYEADDLDPRRTGMSVVVTGLATEVSDPEQVARYQRLLEHWNMAMDTVVAI 135
DB 268 BEGTRVYVY----LKYAKATNLVEVLTVGSEKLDEKGNAR-----KPSSSGAMDNVAIT 318

QY 136 EPEIVTGIRIVAD 148
Db 319 ADEQTNLSLVTAD 331

RESULT 7

ABU28792
ID ABU28792 standard; protein; 654 AA.

XX AC ABU28792;
XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #14319.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Escherichia coli.

XX PN WO20027183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362899P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.

XX DR N-PSDB; ACA32662.

XX PT New antisense nucleic acids, useful for identifying proteins or screening
for homologous nucleic acids required for cellular proliferation to

XX PT isolate candidate molecules for rational drug discovery programs.

XX PS Claim 25; SEQ ID NO 56716; 1766pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of
the 6213 antisense sequences given in the specification where expression
of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 654 AA;

Query Match 11.3%; Score 85.5; DB 6; Length 654;
Best Local Similarity 27.1%; Pred. No. 1.4;
Matches 36; Conservative 20; Mismatches 64; Indels 13; Gaps 4;

QY 20 DRDEAMRLIASVDHGR---VFETRAALPAIRVNVHLVVDGRVIGRTRLTAKV-SVAVRSS 75
Db 208 DLAEILNQLISSEHGKSKOMPALLSAKIVADKXTNSLIISGPEKARQRIITSLKSLDVEES 267

QY 76 ADAGVVVAYEADDLPRRRTGMSVVVTGLATEVSDPEQVARYORLLHPVWNNMAMDTVVAI 135
Db 268 EEGNTRVYV---LKYAKATNLVLEVLTVGSEKLKDEKGNAR-----KPSSSGAMDNVAIT 318

QY 136 EPEIVTGIRIVAD 148
Db 319 ADEQTNLSLVTAD 331

RESULT 8

AAE10129

ID AAE10129 standard; protein; 11096 AA.

XX AC AAE10129;

XX DT 29-NOV-2001 (first entry)

XX DE Streptomyces noursei nystatin gene, NysC.

XX KW Polyketide synthase; PKS; macrolide; nystatin; PKS gene cluster;
XX antifungal; antibiotic; PKS type I.

XX OS Streptomyces noursei.

XX FH Key Location/Qualifiers
XX FT Domain 35..455

FT /label= KS3 domain
FT /note= "Ketosynthase (KS) domain"

FT Domain 546..858

FT /label= AT3 domain
FT /note= "Acyltransferase (AT) domain"

FT Domain 872..1073

FT /label= DH3 domain
FT /note= "Dehydratase (DH) domain"

FT Domain 1381..1628

FT /label= KR3 domain
FT /note= "Ketoreductase (KR) domain"

FT Domain 1662..1735

FT /label= ACP3 domain
FT /note= "Acyl carrier protein (ACP) domain"

FT Domain 1757..2180

FT /label= KS4 domain
FT /note= "Ketosynthase (KS) domain"

FT Domain 2291..2603

FT /label= AT4 domain
FT /note= "Acyltransferase (AT) domain"

FT Domain 2617..2818

FT /label= DH4 domain
FT /note= "Dehydratase (DH) domain"

FT Domain 3124..3371

FT /label= KR4 domain
FT /note= "Ketoreductase (KR) domain"

FT Domain 3407..3480

FT /label= ACP4 domain
FT /note= "Acyl carrier protein (ACP) domain"

FT Domain 3501..3924

FT /label= KS5 domain
FT /note= "Ketosynthase (KS) domain"

FT Domain 4032..4346


```
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI: 2003-029926/02.
DR N-PSDB; ACA37787.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 61841; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 408 AA;
XX
Query Match 10.9%; Score 82.5; DB 6; Length 408;
Best Local Similarity 24.4%; Pred. No. 1.6;
Matches 44; Conservative 23; Mismatches 46; Indels 67; Gaps 10;
QY 3 PGSRASQPSAREVVELDRDEAMLLASVDHGRVFTAAALPAIR-----PVNHL 52
DB 2 PCTRSRGAHRQTVVVSDQ-----RVGGRVSFLPVGMRACAMGVVTVTH 50
QY 53 V-----VDGRVIGRTRLTAKSVAVRSSADAGVW-----AYEADLDLPRRTG---- 96
DB 51 AFQTCHSSGVDGRFLGRFDLAVAVFFAL-----SGFLLWRGHAARDLAPRPTGHYLR 105
QY 97 -----WSVVVTGLATEVSDPDQVARYQLLHP-----WV-NVAMDTVVVAIEPIVTG 142
DB 106 SRVVRIMPAYLVAVVVVILTLPLDAD-----HPSPTVMLANLTLTQIYV--PLTLTG 154
```

```
RESULT 12
ABO79841
ID ABO79841 standard; protein; 1386 AA.
XX
AC ABO79841;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #12016.
XX
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
PF 18-FEB-1999; 99US-00252991.
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA Rubenfield MJ, Nolling J, Deloughery C, Bush D;
PI
XX WPI: 2003-615309/58.
XX N-PSDB; ABD13412.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, of
PT useful as molecular target for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 28587; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
XX Sequence 1386 AA;
XX
Query Match 10.9%; Score 82.5; DB 7; Length 1386;
Best Local Similarity 29.7%; Pred. No. 8.4;
Matches 35; Conservative 14; Mismatches 44; Indels 25; Gaps 4;
QY 3 PGSRASQPSAREVVELDRDE-----AMLLASVDHGRV-----VFT 39
DB 154 PGGRRGTPPGASADSELHRRHLELPLGLAHRHVOSIGTRVVGAVEQVLHVLELHVA 213
QY 40 RAALPAIRPNVNLVVDGRVIGRTRLTAKSVAVRSSADAGVWVAYE-ADDDLPRRTG 96
DB 214 EAGVVAGEGIGH-GVGQGVGIGRVARALADVARAAQAQAVAAEELAEFVDRRQRG 270
XX
XX RESULT 13
ABU47143
ID ABU47143 standard; protein; 454 AA.
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 01:42:44 ; Search time 12.4799 Seconds
(without alignments)
909.197 Million cell updates/sec

Title: US-10-617-038-2
Perfect score: 755
Sequence: 1 MSPGSRASQSAREVVELD.....VAIEPIVGTIRIVADSRTP 152

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
 - 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
 - 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
 - 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
 - 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
 - 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90.5	12.0	655	4	US-09-902-540-10005
2	85	11.3	269	4	US-09-902-540-16003
3	82.5	10.9	1386	4	US-09-902-540-16003
4	81.5	10.8	505	4	US-09-489-039A-14050
5	80.5	10.7	204	4	US-09-252-991A-29225
6	79.5	10.5	1490	4	US-09-252-991A-28442
7	79	10.5	547	4	US-09-902-540-13219
8	77.5	10.3	163	4	US-09-540-236-2257
9	77.5	10.3	688	4	US-09-902-540-10297
10	77	10.2	1144	4	US-09-252-991A-26588
11	77	10.2	1612	3	US-08-545-860D-48
12	77	10.2	1612	5	PCT-US94-04496-48
13	76.5	10.1	328	4	US-09-724-797-8
14	75.5	10.0	585	4	US-09-252-991A-30045
15	75.5	10.0	1151	4	US-09-252-991A-21328
16	74.5	9.9	195	4	US-09-252-991A-26313
17	74.5	9.9	263	4	US-09-252-991A-25669
18	74.5	9.9	495	4	US-09-252-991A-26109
19	74	9.8	232	3	US-09-555-270A-13
20	74	9.8	270	3	US-09-551-941-15
21	74	9.8	270	4	US-09-955-597-15
22	74	9.8	425	4	US-09-252-991A-33074
23	74	9.8	536	4	US-09-902-540-10280
24	74	9.8	4150	3	US-09-428-517-2
25	73.5	9.7	488	4	US-09-902-540-13224
26	73.5	9.7	764	4	US-09-252-991A-31816
27	73.5	9.7	3025	6	5223423-3

28 73.5 9.7 3025 6 5223423-3 Patent No. 5223423
29 73.5 9.7 7831 4 US-09-902-540-12902 Sequence 12902, A
30 73 303 4 US-09-902-540-14250 Sequence 14250, A
31 73 902 4 US-09-107-532A-3837 Sequence 3837, Ap
32 72.5 9.6 345 4 US-09-902-540-11021 Sequence 11021, A
33 72.5 9.6 346 4 US-09-252-991A-27167 Sequence 27167, A
34 72.5 9.6 548 2 US-08-467-822-31 Sequence 31, Appl
35 72.5 9.6 548 3 US-08-432-697-31 Sequence 31, Appl
36 72.5 9.6 548 3 US-08-466-248-31 Sequence 31, Appl
37 72 9.5 437 2 US-09-031-059-1 Sequence 3, Appl
38 72 9.5 437 2 US-09-031-059-3 Sequence 3, Appl
39 72 9.5 711 4 US-09-252-991A-28660 Sequence 28660, A
40 71.5 9.5 635 4 US-09-252-991A-30924 Sequence 30924, A
41 71.5 9.5 671 4 US-09-252-991A-19016 Sequence 19016, A
42 71.5 9.5 4551 3 US-09-320-878-1 Sequence 1, Appl
43 71.5 9.5 4551 4 US-09-141-908-2 Sequence 2, Appl
44 71.5 9.5 4551 4 US-09-657-440-1 Sequence 1, Appl
45 71.5 9.5 4613 3 US-09-105-537-31 Sequence 31, Appl

ALIGNMENTS

RESULT 1
US-09-902-540-10005
; Sequence 10005, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10005
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10005

Query Match 12.0%; Score 90.5; DB 4; Length 655;
Best Local Similarity 25.3%; Pred. No. 0.029;
Matches 37; Conservative 25; Mismatches 69; Indels 15; Gaps 4;
QY 11 QSAREVVELDRDEAMRLASVDHGR-----VVFTRAAALPAIRPVNHLVVDGRVIGRTRL 64
Db 239 QQARAIAEALSQAREALAAEQARAEALASQATLEAEQOTR--ADTGAALQOTRS 296
QY 65 TAKSVAVRVSSADAGVVVAYEADDLDPRRTGWSVVVTGLATEVSDPEQVARYQRLHLHPW 124
Db 297 ALESEQQGRAEALAAQVRAVLEAEQRTTDAESALQTLRTETSDQ--ARQERL---- 351
QY 125 VNAMDVTVAIEPIVGTIRIVADSR 150
Db 352 --AAAEVMAEQERALEAQLAAEAR 375

RESULT 2
US-09-902-540-16003
; Sequence 16003, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B

```
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2003-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16003
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
; US-09-902-540-16003

Query Match
Best Local Similarity 11.3%; Score 85; DB 4; Length 269;
Matches 46; Conservative 20; Mismatches 42; Indels 60; Gaps 8;

QY 13 AREVVELDRDEAMRLLA-----SVDHGRVVFTRA-----ALPAI 46
Db 72 AAGLVALVRDPAIILLIGGLFAVAMSIOPLSLSAFQVLLTPALVLLAELQSGDWELAGV 131
QY 47 RPNHLVVDGRVIGRT-----RLTAKVSVAVRSSADAGVVVYAEADDLP--- 91
Db 132 RIVNTLL--GGUIALTVTLLWPSPEHLRLPEQVAFALRADREYLMVAVAAHSDSEPAVR 189
QY 92 --RRRTGWSVVVTGLATEVSDPEQVARYORLLHPVWMMAMDTVVAIEP 137
Db 190 EARRKIGLAL-----LAAEAS-----FORLLSEWTGPXKD-----LEP 222

RESULT 3
US-09-252-991A-28587
; Sequence 28587, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28587
; LENGTH: 1386
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1294)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-28587

Query Match
Best Local Similarity 10.9%; Score 82.5; DB 4; Length 1386;
Matches 35; Conservative 14; Mismatches 44; Indels 25; Gaps 4;

QY 3 PGSRRASPOSAREVVELDRDE-----AMRLLASVDHGRV-----VFT 39
Db 154 PGGRRCTPGCASDSELHRRRHLELPGLAHRHVQSIGTGRVVGAVEQVHLHVELLHAYSA 213
QY 40 RAALPAIRPNHLVVDGRVIGRTLRITAKVSVAVRSSADAGVVVYAE-ADDLPDRRTG 96
Db 214 EAGVAGEGIGH-GVGGQVEGIGRVARALADVARAAQAQVAAABELAEFVDRRQRGG 270

RESULT 4
US-09-489-039A-14050
; Sequence 14050, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 14050
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-14050

Query Match
Best Local Similarity 10.8%; Score 81.5; DB 4; Length 505;
Matches 38; Conservative 20; Mismatches 50; Indels 49; Gaps 8;

QY 8 ASPOSAREVVELDRDEAMRLLASVDHGRVVFTRAALP-----AIRP---VNHLVVD 55
Db 337 ADAQAEKLANSPKDRANLMI-VLMNRNDIGRVAVFGSVRPDLFVVVEPPPAVHLV-- 393
QY 56 GRVIGRTLRITAKVSVAVRSSADAGVVV-----AYE-ADDLPDRRTG--- 97
Db 394 STITARLPMTLHASDLLRAAFPGGISITGAPKVRAMEIIDLEPQRRNAWCGSIGVLSYCG 453
QY 98 ----SVVVTGLAT-----EVSDPEQVARYQ 118
Db 454 NMDTSITIRTLTAWQGLYCSAGGGIVADSEAAEYQ 490

RESULT 5
US-09-252-991A-29225
; Sequence 29225, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29225
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-29225

Query Match
Best Local Similarity 10.7%; Score 80.5; DB 4; Length 204;
Matches 42; Conservative 13; Mismatches 51; Indels 49; Gaps 7;

QY 5 SRRASPOSAREVVELDRDEAMRLLASVDHGRVVFTRAALPAI-RPVNHLVVDGRVIGTR 63
Db 37 SRRTSGRGTGRPVRRASRSGPARGLP--EGR---RRRLPGLPRP-----LPGALPGRTD 85
QY 64 LTAKVSVAVRSSADAGVVVYAEADDLD-----PRRITGWSVVVTCCLATEV 108
Db 86 HPAPRRRVSATAGAAGVQQHDPPEAAAGRLTGGGRPPCHPQRR-----LPGLQTRA 140
QY 109 SDPEQVARYQRL-----LHPVWN 126
Db 141 ERPVSQARPQQLPDLRRPVAAQLPQGGPDPAEPWLN 175

RESULT 6
US-09-252-991A-28442
; Sequence 28442, Application US/09252991A
```

; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28442
; LENGTH: 1490
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28442

Query Match 10.5%; Score 79.5; DB 4; Length 1490;
Best Local Similarity 22.4%; Pred. No. 2.2;
Matches 44; Conservative 20; Mismatches 69; Indels 63; Gaps 7;

QY 14 REVVELDRDEAMRLASVDHGRVVFTRAAALPAIRPVNHLVVDGRVIGRTRLTAKV----- 68
Db 284 RVVLHVQQTTHAPOOLGALHRRARRVGRGALFAVQRMQHRV--GQLLGRDLRGEKIDEGRR 341
QY 69 SVAVRSSADAGVV-----VAYEADD----- 88
Db 342 AVVARIGQHLAIGDHHGRQVALQADQLRRAEAVEIGHAPTHQHHVVATALGHR 401
QY 89 LDPRRRTGMSVVVTGLATEVSDP--EQVARY-----QRLHHPWVNMAM-----DTV 132
Db 402 LDPHRLAAADAAGAPAGADRAFQLAGHVVRHQQPPQFAGGAGRTQRLVVGDSVE 461
QY 133 VAIEPEIVTGTIRIVAD 148
Db 462 VQVEPEPATAARRALD 477

RESULT 7
US-09-902-540-13219
; Sequence 13219, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13219
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13219

Query Match 10.5%; Score 79; DB 4; Length 547;
Best Local Similarity 28.9%; Pred. No. 0.59;
Matches 39; Conservative 13; Mismatches 55; Indels 28; Gaps 6;

QY 16 VVELDRDEAMRLASVDHGRVVFTRAAALPAIRPVNHLVVDG-----RVIGRTRLTAKVSV 70
Db 300 VVSEDLGHKFETLTLDLG-----RAKRVTVDKDNTVVDGVGTAKAIEGRIKL----- 348
QY 71 AVRSSADAGVWAYEADDLDPRTT---GWSVVVTGLATEVSDPEQVARYQRLHHPWVNM 127
Db 349 -TQTQIDS-VTSDYDEKLEQLERLAKLVGGVAVINVGAAITETEMKEKARVEDALHA----- 402

QY 128 AMDTVVAIEPEIVTG 142
Db 403 ---TRAAVEEGIVPG 414

RESULT 8
US-09-540-236-2257
; Sequence 2257, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRH
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2257
; LENGTH: 163
; TYPE: PRT
; ORGANISM: M. catarrhalis
US-09-540-236-2257

Query Match 10.3%; Score 77.5; DB 4; Length 163;
Best Local Similarity 25.2%; Pred. No. 0.16;
Matches 35; Conservative 25; Mismatches 52; Indels 27; Gaps 7;

QY 6 RRASPOSAREVVELDRDEAMRL--LASVDHGRVVFTRAAALPAIRPVNH---LVVDGRVIG 60
Db 28 KRIQPMISVEMVELPPAKRSKTPSPAEDIKYKAIETTLNKAHPKERLWLLEVKGMLS 87
QY 61 RTRLTAKVSVAVRSSADAGVWAYEADDLDPRTTGT---WSVVVTGLATEVSDP----- 111
Db 88 TEHLSDKLSVAMQDCADIALVIG-GADGVSPEILQAADFKWSL-----SELTLPHPLVR 140
QY 112 -----EQVARYQRL--HPW 124
Db 141 IIVIEQLYRAMSLINHHPY 159

RESULT 9
US-09-902-540-10297
; Sequence 10297, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10297
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10297

Query Match 10.3%; Score 77.5; DB 4; Length 688;
Best Local Similarity 28.8%; Pred. No. 1.3;
Matches 30; Conservative 13; Mismatches 40; Indels 21; Gaps 3;

QY 48 PVNHLVVDGRVIGRTRLTAKVSVAVRSSADAGVWAYEADDLDPRTTGTGMSVVVTGLATE 107
Db 563 PVLVVPADGEVVDSTPLFA-----GAATAGATVVIEVDGIE-----VATV 602
QY 108 VSDPEQVARYQRLHHPWVNMAMDTVVVAIEPEIVTGTIRIVADSRT 151

Db 603 TADDTGAFRYQVPRESALVSGSHVAAQE-RLVTSSRVPARSRT 645

RESULT 10
US-09-252-991A-26588
; Sequence 26588, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26588
; LENGTH: 1144
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26588

Query Match 10.2%; Score 77; DB 4; Length 1144;
Best Local Similarity 27.1%; Pred. No. 3.1;
Matches 45; Conservative 17; Mismatches 72; Indels 32; Gaps 5;

Qy 3 PGSRASQSAAREVVELDRDEAMRLASVDHGRVVFTRAAALPAIRPVNHLV-----VDG 56
Db 16 PHSRRERPRGREVQGPPLPDPVAVDTGRQERLADHVVAQDPWVLADRTVDL 75
Qy 57 RVIG---RTRLTAKSVAVRSADAG-----VVVAYEADDLDPRRRTGWSVVVTGL 107
Db 76 AVMLRLVLEADLAADAGVLDRSQVDAGQPAHGVDVVVATDGAFRHGRNDASLDLVLADQAAD 135
Qy 108 VSDPEQVARYQRLHPVNMAMDTVVAIEPEIVTGIRI--VADSRT 151
Db 136 AAAP-----AAVDTGTAARRODRTGVAADQAANRRT 166

RESULT 11
US-08-545-860D-48
; Sequence 48, Application US/08545860D
; Patent No. 6040140
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Canaan, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
; TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
; TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1 Region
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESS: No. 6040140ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,860D
; FILING DATE: 07-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04496
; FILING DATE: 22-APR-1994

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10930
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/327,392
; FILING DATE: 19-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/320,559
; FILING DATE: 11-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/062,443
; FILING DATE: 14-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/971,094
; FILING DATE: 30-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,839
; FILING DATE: 27-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/805,093
; FILING DATE: 11-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca Esq., Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1262
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1612 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-545-860D-48

Query Match 10.2%; Score 77; DB 3; Length 1612;
Best Local Similarity 24.3%; Pred. No. 5.1;
Matches 42; Conservative 22; Mismatches 69; Indels 40; Gaps 8;

Qy 1 MSPGSRASQSAAREVVELDRDEA-MELLAS--VDHGRVVF-----TRAAPL 44
Db 301 LPQAQHSDEKGAKEII-LDDDECPLOIFREWPSDKGILVQLKRRPPDHIPKTKKHLE 359
Qy 45 AIREVNHVVDGRVIGRTRLTAKSVAVRSADAGVVVAVYEAADDLPPRRRTGWSVVVTGL 104
Db 360 GKTPGKERADGSVYGSTLPPEKLPYLVELSPDG-----SDSRDKPKLYRLQLSVTEV 412
Qy 105 ATEVSDPEQVARYORLHPVNM-----MAMDTVVAIEPE-----IVTGIRI 145
Db 413 GTEKLDNDSI----QLFGPGIQQPHHCDLTNNMDGVVTVTPRSMDAETVEGORI 461

RESULT 12
PCT-US94-04496-48
; Sequence 48, Application PC/TUS9404496
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Canaan, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
; TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
; TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESS: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA

; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04496
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca Esq., Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1242
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1612 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; PCT-US94-04496-48

Query Match 10.2%; Score 77; DB 5; Length 1612;
Best Local Similarity 24.3%; Pred. No. 5.1;
Matches 42; Conservative 22; Mismatches 69; Indels 40; Gaps 8;

QY 1 MSPGSRASPOSAREVELDRDEA-MRLLAS--VDHGRVVF-----TRAALP 44
DB 301 LPFGAQSDEKAKEII-LDDDECPQIREFWPSDKGILVQLKRRPPDHPKPKTKHLE 359
QY 45 ATRPNHLVVDGRVIGRTRLTAKVSVAVRSSADAGVVVAYEADDLPDRRTGWSVVVTGL 104
DB 360 GKTPKGRADGSVYSTLPPEKLPYLVELSPDG-----SDSRDKPLYRLQLSVTEV 412
QY 105 ATEVSPQEQVARYQLLHPVWN-----MAMDTVVAIEPE-----IVTGIRI 145
DB 413 GTEKLDNSI-----QLFGFGIQQPHCHDLTNMGCVVTVTPRSMDAETVVEGQRI 461

RESULT 13
US-09-724-797-8
; Sequence 8, Application US/09724797
; Patent No. 6733998
; GENERAL INFORMATION:
; APPLICANT: Jon S. THORSON
; TITLE OF INVENTION: MICROMONOSPORA ECHINOSPORA GENES
; TITLE OF INVENTION: ENCODING FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: CALICHEAMICIN AND SELF-RESISTANCE THERETO
; FILE REFERENCE: 2653-40
; CURRENT APPLICATION NUMBER: US/09/724,797
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 60/111,325
; PRIOR FILING DATE: 1998-12-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Bacteria
US-09-724-797-8

Query Match 10.1%; Score 76.5; DB 4; Length 328;
Best Local Similarity 30.3%; Pred. No. 0.57;
Matches 40; Conservative 12; Mismatches 31; Indels 49; Gaps 8;

QY 10 PQSAREVELDR-----DEAMRL-----LASVDHGRVVF--TRAALPAIRP 48
;
;
;
;

DB 172 PRSRSEWEMIRALVRDGGTVLLTTQYLDADHLADELTLIDHGRIVACGTPELKASRA 231
QY 49 VNHLVVDGRVIGRTRLTAKVSVAVRSSADAGVVVA-----YEAD-DLDPRRRTGWSVVVTG 103
DB 232 AG--VLDVLRDPER-----RADAGALLAKAVCAAADLSDPAR----- 268
QY 104 LATEVSDPEQVA 115
DB 269 LSVRVTPDRAA 280

RESULT 14
US-09-252-991A-30045
; Sequence 30045, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30045
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30045

Query Match 10.0%; Score 75.5; DB 4; Length 585;
Best Local Similarity 25.4%; Pred. No. 1.8;
Matches 48; Conservative 22; Mismatches 64; Indels 55; Gaps 10;

QY 2 SPG-SRRASPOSAREVELDRDEAMRLLASVDHGRVVFTRAAALPAI----- 46
DB 56 APGLGHRAELRCARLV-----GRALRRDAPEVHSRLAMVEVVAGQVLHEGHPQDVDOQ 110
QY 47 ---RPNHLVVDGR--VIGRTRLTAKVSVAVRS---SADAGVVVAYE-----ADDLPDRR 94
DB 111 HAQAPANERQGEQALLVGRVGAABEASPAQRAVVQEAQAVLVAGEPEVLGVDLPDRV 170
QY 95 TGWSVVVTGLATEVSDPEQVARYQL-----LHPW-----VNMAMDTVVA---IEPEIVT 141
DB 171 LG-----EDEDREVEQERERLQVAAGVHQAEOQSPVERPADALVARGGDEADVET 220
QY 142 GIRIVADSR 150
DB 221 ELQVLEDPDR 229

RESULT 15
US-09-252-991A-21328
; Sequence 21328, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21328
; LENGTH: 1151
; TYPE: PRT
;

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21328

Query March	10.0%;	Score 75.5;	DB 4;	Length 1151;
Best Local Similarity	33.6%;	Pred. No. 4.8;		
Matches 43;	Conservative 7;	Mismatches 49;	Indels 29;	Gaps 7;
Qy	15	EWVEL----	DRDEAMR-----	LLASVDHGVRVFTRAALPAIRP-----VNHLLVVDG 56
Db	950	EFVELAVDPGADENRAHLV	EHROVLALALADHRR----	QOHLAAFRQGGDLVDHL-ADG 1005
Qy	57	RVIGRTRLTAKVSVAVRSSADAGVVVAYE	ADDLDPRRRITGMSVVVTGLATEVSDPEQV--	114
Db	1006	-----LRLQRDVVVRAARGTDG	IKQIKQVVVDLGDRTHGGAWVGSGLLFDGNNRRRPPD	1060
Qy	115	ARYORLLH 122		
Db	1061	GVHVRLLH 1068		

Search completed: October 15, 2005, 03:24:05
Job time : 14.4799 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:14:10 ; Search time 36.3491 Seconds
(without alignments)
1742.576 Million cell updates/sec

Title: US-10-617-038-2
Perfect score: 755
Sequence: 1 MSPGRRASPOQAREVVELD.....VAIEPIVTGIRIVADSRTP 152

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA.*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
 - 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
 - 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	755	100.0	152	15	US-10-617-038-2
2	205	27.2	208	14	Sequence 2, Appli
3	98.5	13.0	164	14	Sequence 8810, Ap
4	85.5	11.3	192	15	Sequence 11908, A
5	85.5	11.3	654	9	Sequence 19770, A
6	85.5	11.3	654	15	Sequence 10361, A
7	84.5	11.2	11088	16	Sequence 56716, A
8	84.5	11.2	11096	17	Sequence 7, Appli
9	83	11.0	304	15	Sequence 20557, A
10	82.5	10.9	110	15	Sequence 8090, Ap
11	82.5	10.9	408	15	Sequence 38, Appli
					Sequence 61841, A

12	82.5	10.9	852	16	US-10-437-963-164426	Sequence 164426,
13	82	10.9	454	15	US-10-282-122A-75067	Sequence 75067, A
14	82	10.9	454	15	US-10-282-122A-75696	Sequence 75696, A
15	81.5	10.8	451	15	US-10-282-122A-60087	Sequence 60087, A
16	81.5	10.8	577	15	US-10-282-122A-50518	Sequence 50518, A
17	81	10.7	445	15	US-10-282-122A-72978	Sequence 72978, A
18	80.5	10.7	329	15	US-10-104-047-3709	Sequence 3709, Ap
19	80.5	10.7	3931	15	US-10-120-801-18	Sequence 18, Appl
20	80	10.6	518	15	US-10-369-493-20436	Sequence 20436, A
21	80	10.6	739	14	US-10-156-761-13097	Sequence 13097, A
22	80	10.6	1180	15	US-10-369-493-8316	Sequence 8316, Ap
23	79	10.5	448	15	US-10-282-122A-62698	Sequence 62698, A
24	79	10.5	448	15	US-10-282-122A-64857	Sequence 64857, A
25	79	10.5	452	14	US-10-156-761-12413	Sequence 12413, A
26	79	10.5	476	16	US-10-739-930-8412	Sequence 8412, A
27	78.5	10.4	288	14	US-10-156-761-11595	Sequence 11595, A
28	78.5	10.4	326	15	US-10-424-599-192288	Sequence 192288, A
29	78.5	10.4	919	15	US-10-369-493-5367	Sequence 5367, Ap
30	78.5	10.4	984	17	US-10-211-028-25	Sequence 25, Appl
31	78	10.3	288	15	US-10-461-194-140	Sequence 140, App
32	78	10.3	813	15	US-10-369-493-3604	Sequence 3604, Ap
33	77.5	10.3	160	15	US-10-282-122A-63336	Sequence 63336, A
34	77.5	10.3	235	15	US-10-425-114-52734	Sequence 52734, A
35	77	10.2	288	15	US-10-461-194-108	Sequence 108, App
36	77	10.2	366	15	US-10-369-493-13763	Sequence 13763, A
37	77	10.2	682	14	US-10-156-761-11449	Sequence 11449, A
38	77	10.2	1612	20	US-11-037-713-1	Sequence 1, Appli
39	77	10.2	9234	10	US-09-942-025-13	Sequence 13, Appl
40	76.5	10.1	568	15	US-10-369-493-6846	Sequence 6846, Ap
41	76.5	10.1	741	15	US-10-282-122A-49230	Sequence 49230, A
42	76	10.1	258	15	US-10-084-846A-90	Sequence 90, Appl
43	76	10.1	376	14	US-10-329-079-65	Sequence 65, Appl
44	76	10.1	527	15	US-10-369-493-10548	Sequence 10548, A
45	76	10.1	571	14	US-10-156-761-10211	Sequence 10211, A

ALIGNMENTS

RESULT 1
US-10-617-038-2
; Sequence 2, Application US/10617038
; Publication No. US20040057963A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Peter
; APPLICANT: Rosenkrands, Ida
; APPLICANT: Stryhn, Anette
; TITLE OF INVENTION: Therapeutic TB Vaccine
; FILE REFERENCE: SS15AUSA
; CURRENT APPLICATION NUMBER: US/10/617,038
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: DK PA 2002 01098
; PRIOR FILING DATE: 2002-07-13
; PRIOR APPLICATION NUMBER: US 60/401,725
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-617-038-2

Query Match	100.0%	Score 755;	DB 15;	Length 152;
Best Local Similarity	100.0%	Pred. No. 1.4e-76;		
Matches 152;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Oy	1	MSPGRRASPOQAREVVELDRDEAMRLLASVDHGRVVFTRAAALPAIRPVNHLVVDGRVIG	60	
Db	1	MSPGRRASPOQAREVVELDRDEAMRLLASVDHGRVVFTRAAALPAIRPVNHLVVDGRVIG	60	
Qy	61	RTLTAKVSVAVRSSADAGVVVAYEADLLDPRRTGWSVVVTGLATEVSDPEQVARYORL	120	

Db 61 RTRLTAKVSVVRSSADAGVVAEADDDLPRTTGWSSVVVTGLATEVSDPEQVARYQRL 120

QY 121 LHPWNNAMDTVVAIEPEIVTGTIRIVADSRTP 152

Db 121 LHPWNNAMDTVVAIEPEIVTGTIRIVADSRTP 152

RESULT 2

US-10-156-761-8810

Sequence 8810, Application US/10156761

Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI

APPLICANT: SHIBA, TADAYOSHI

APPLICANT: SAKAKI, YOSHIYUKI

APPLICANT: HATTORI, MASAHIRA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262

CURRENT APPLICATION NUMBER: US/10/156,761

CURRENT FILING DATE: 2002-05-29

PRIOR APPLICATION NUMBER: JP 2001-204089

PRIOR FILING DATE: 2001-05-30

PRIOR APPLICATION NUMBER: JP 2001-272697

PRIOR FILING DATE: 2001-08-02

NUMBER OF SEQ ID NOS: 15109

SEQ ID NO 8810

LENGTH: 208

TYPE: PRT

ORGANISM: Streptomyces avermitilis

US-10-156-761-8810

Query Match 27.2%; Score 205; DB 14; Length 208;

Best Local Similarity 37.2%; Pred. No. 1.6e-14;

Matches 55; Conservative 22; Mismatches 61; Indels 10; Gaps 3;

QY 1 MSPGSRASPSQSAREWELDRDEANRLLASVDHGRVTFTRALPAIRPVNHLVDGVRIG 60

Db 67 LPPGVGRAAHF--ELLESDSECRARLSTHGVGLAVDTPTGPVIVPLNYSVWDGVVCF 124

QY 61 RTRLTAKVSVVRSSADAGVVAEADDDLPRTTGWSSVVVTGLATEVSDPEQVARYQRL 120

Db 125 RTAADSE-----PAASAGSRVAFEDVDHIDEALSQGSNVLVRGLARLVTDPTLRLLAEL 178

QY 121 LH--PWNNAMDTVVAIEPEIVTGTIRIV 146

Db 179 AYSGPWAGGERDTWVCVDPVGTGRRIV 206

RESULT 3

US-10-156-761-11908

Sequence 11908, Application US/10156761

Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI

APPLICANT: SHIBA, TADAYOSHI

APPLICANT: SAKAKI, YOSHIYUKI

APPLICANT: HATTORI, MASAHIRA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262

CURRENT APPLICATION NUMBER: US/10/156,761

CURRENT FILING DATE: 2002-05-29

PRIOR APPLICATION NUMBER: JP 2001-204089

PRIOR FILING DATE: 2001-05-30

PRIOR APPLICATION NUMBER: JP 2001-272697

PRIOR FILING DATE: 2001-08-02

NUMBER OF SEQ ID NOS: 15109

SEQ ID NO 11908

Query Match 11.3%; Score 85.5; DB 15; Length 192;

Best Local Similarity 31.4%; Pred. No. 0.41;

Matches 33; Conservative 15; Mismatches 24; Indels 33; Gaps 6;

QY 65 TAKVSVAVRSSADAGV-----VVAYEADDDLPRTT-----TWSSVVVTGL 104

Db 3 TSKEPVVITGASAGIGRATVRAFAKRGAYTGLVARSRDGLEAAARQVEAAAGAKALV--L 60

QY 105 ATEVSDPEQV----ARYORLLHP---WVNMAVMDTVVA-----IEPE 138

Db 61 PTDVSDPEQVEAAAAAAVEKEFGPIDIWNVDAMASILSPFTEIKPE 105

RESULT 5

US-09-815-242-10361

Sequence 10361, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

; LENGTH: 164

; TYPE: PRT

; ORGANISM: Streptomyces avermitilis

US-10-156-761-11908

Query Match 13.0%; Score 98.5; DB 14; Length 164;

Best Local Similarity 31.1%; Pred. No. 0.011; Indels 29; Gaps 7;

Matches 41; Conservative 14; Mismatches 14; Indels 29; Gaps 7;

QY 24 AMRLASVDHGRVTFTRALPAIRPVNHLVDGVRIGRTRLTAKVSVAVRS-----SADAG 79

Db 8 AVELIGRADHGRVATSMRALPFLVCVRHIVLAGRVLLRLH-----RCMGYHRACVG 58

QY 80 VVAYEADDDLPRTTG-----WSVVVTGLATEVSDPEQVARYQRLLHPWNNMA-----MD 130

Db 59 SVVAYGSDNLG---RPGGGNLSAQIVGRCEAI---EPTAAQIALFGAPRVDGRPFD 112

QY 131 TV-VAIEPEIVT 141

Db 113 PVYLRIEPOLGT 124

RESULT 4

US-10-369-493-19770

Sequence 19770, Application US/10369493

Publication No. US20030233675A1

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Goldman, Barry S.

APPLICANT: Chen, Xianfeng

TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

FILE REFERENCE: 38-10(52052)B

CURRENT APPLICATION NUMBER: US/10/369,493

CURRENT FILING DATE: 2003-02-28

PRIOR APPLICATION NUMBER: US 60/360,039

PRIOR FILING DATE: 2002-02-21

NUMBER OF SEQ ID NOS: 47374

SEQ ID NO 19770

LENGTH: 192

TYPE: PRT

ORGANISM: No. US20030233675A1toc punctiforme

US-10-369-493-19770

Query Match 11.3%; Score 85.5; DB 15; Length 192;

Best Local Similarity 31.4%; Pred. No. 0.41;

Matches 33; Conservative 15; Mismatches 24; Indels 33; Gaps 6;

QY 65 TAKVSVAVRSSADAGV-----VVAYEADDDLPRTT-----TWSSVVVTGL 104

Db 3 TSKEPVVITGASAGIGRATVRAFAKRGAYTGLVARSRDGLEAAARQVEAAAGAKALV--L 60

QY 105 ATEVSDPEQV----ARYORLLHP---WVNMAVMDTVVA-----IEPE 138

Db 61 PTDVSDPEQVEAAAAAAVEKEFGPIDIWNVDAMASILSPFTEIKPE 105

RESULT 5

US-09-815-242-10361

Sequence 10361, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

Query Match 11.3%; Score 85.5; DB 15; Length 654;
Best Local Similarity 27.1%; Pred. No. 2.1;
Matches 36; Conservative 20; Mismatches 64; Indels 13; Gaps 4;
QV 20 DRDEAMRLLASVDHGR---VFTRTALPAIRPVNHLWDGVRIGTRLTAKV-SVAVRSS 75

Query Match	11.3%;	Score	85.5;	DB	15;	Length	654;
Best Local Similarity	27.1%;	Pred.	No. 2.1;				
Matches	36;	Conservative	20;	Mismatches	64;	Indels	13; Gaps 4;
Qy	20	DRDEANRLNASVDHGR--VVFTRPALPAIRPNVHLVVGDGVRIGRTLTAKV-SVAVRSS	75				
Dd	208	DUAETLNLQIISEHGKSQMPALLSAKIVADKRTNSLIISGPEKARORITSLLSLDVEES	267				
Qy	76	ADAGVVVYAEADDLPERRRTGMSVVVTGLATEVSDDPEQARYQRLLHPWVNMMAMDTVAI	135				
Dd	268	EENGTRVYY----LKAYAKATNLVELVTGVSEKLKDEKNAR-----KPSSSGAMDNVAIT	318				
Qy	136	EPEIVTGIIRIVAD	148				
Dd	319	ADEQTNSLVITAD	331				

```

RESULT 7
US-10-203-295-7
; Sequence 7, Application US/10203295
; Publication No. US20040115762A1
; GENERAL INFORMATION:
; APPLICANT: Zorchev, Sergey Borisovich
; APPLICANT: Sekurova, Olga Nikalayivna
; APPLICANT: Fjaervik, Epsen
; APPLICANT: Broutaset, Trygve
; APPLICANT: Strom, Arne Reidar
; APPLICANT: Valla, Svein
; APPLICANT: Ellingsen, Trond Erling
; APPLICANT: Sletta, Havard
; APPLICANT: Gulliksen, Ole-Martin
; TITLE OF INVENTION: Novel genes encoding a nystatin polyketide synthase and their
; TITLE OF INVENTION: manipulation and utility
; FILE REFERENCE: 1181-265
; CURRENT APPLICATION NUMBER: US/10/203,295
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: PCT/GB 01/00509
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: GB 0002840.7
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: GB 0008786.6
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: GB 0009387.2
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 11089
; TYPE: PRT
; ORGANISM: Streptomyces noursei ATCC 11455
US-10-203-295-7

```

Query Match 11.2%: Score 84.5: DB 16: Length 11088:


```

; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61841
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Mycobacterium avium
US-10-282-122A-61841

Query Match      10.9%; Score 82.5; DB 15; Length 408;
Best Local Similarity 24.4%; Pred. No. 2.4;
Matches 44; Conservative 23; Mismatches 46; Indels 67; Gaps 10;

QY 3 PGSRASPOSAREVVELDRDEAMRLLASVDHGRVWFTRAALPAIR-----PVNHL 52
Db |||: : : : : |||: : : : : |||: : : : : |||: : : : :
QY 2 PGCTSPGRAHRTVVVSDQ-----RVGVRGFLFAVEGRACAAAGVVVVTHV 50
Db |||: : : : : |||: : : : : |||: : : : : |||: : : : :
QY 53 V-----VDGRVIGRTRLTAKVSVAVRSSADAGVV-----AYEADDLPRTTG----- 96
Db |||: : : : : |||: : : : : |||: : : : : |||: : : : :
QY 51 AFQTHSSGVDRGLFGRFDLAVAVFFAL-----SGFLWRGHAAAAARDLAPRPTGHYLR 105
Db |||: : : : : |||: : : : : |||: : : : : |||: : : : :
QY 97 -----WSVVVTGLATEVSDEQVARYQRLHP-----WV-NMAMDTVVAIEPIVVG 142
Db |||: : : : : |||: : : : : |||: : : : : |||: : : : :
QY 106 SRVVRIMPAYLAVAVVVVILTLPPAD-----HPSPTVWLANLTLTIQYV--PLTLTG 154
Db |||: : : : : |||: : : : : |||: : : : : |||: : : : :

RESULT 12
US-10-437-963-164426
; Sequence 164426, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 164426
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_63326C.1.p
US-10-437-963-164426
```

```

Query Match      10.9%; Score 82.5; DB 16; Length 852;
Best Local Similarity 33.1%; Pred. No. 6.4;
Matches 42; Conservative 13; Mismatches 27; Indels 45; Gaps 8;

QY 12 SARE---VVELDRDEAM-RLLASVDH--GRVVFTRAALPAIRPVNHLVVVDGRVIGTRLT 65
Db |||: : : : : |||: : : : : |||: : : : : |||: : : : :
QY 708 SURETFCLAQADGGKAVERLVACLDHLDGRVV--EALAAAL-----STLVCDG----- 753
Db |||: : : : : |||: : : : : |||: : : : : |||: : : : :
QY 66 AKSVSAVRSSADAGVVVAYEADDLP-----RRRTGWSW-----VTGLATEV 108
Db |||: : : : : |||: : : : : |||: : : : : |||: : : : :
QY 109 SDPEQVA 115
Db : : : : :
QY 806 AADQTV 812
Db : : : : :

RESULT 13
US-10-282-122A-75067
; Sequence 75067, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75067
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Salmonella typhimurium
US-10-282-122A-75067

Query Match      10.9%; Score 82; DB 15; Length 454;
Best Local Similarity 25.7%; Pred. No. 3.2;
Matches 44; Conservative 22; Mismatches 53; Indels 52; Gaps 10;

QY 7 RASPOSAREVVELDRDEAMRLLASVDHGRVWFTRAALP-----AIRP-----VNHLV- 53
```


Search completed: October 15, 2005, 04:14:55
Job time : 37.3491 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:59:30 ; Search time 8.48147 Seconds
(without alignments)
1724.342 Million cell updates/sec

Title: US-10-617-038-2

Perfect score: 755

Sequence: 1 MSPGRRASPOQAREVVELD.....VAIEPIVTGIRIVADSRTP 152

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79;.*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	755	100.0	152	2 A70850	hypothetical prote
2	225.5	29.9	146	2 T37028	hypothetical prote
3	223.5	29.6	137	2 T37043	hypothetical prote
4	178	23.6	323	2 T36956	probable DNA-bindi
5	131	17.4	233	2 T36942	probable DNA-bindi
6	103	13.6	135	2 T36500	hypothetical prote
7	90.5	12.0	437	2 T46573	UDPGlucose 6-dehyd
8	85.5	11.3	654	2 H65125	probable general s
9	84	11.1	380	2 JN0823	schA protein - Str
10	82	10.9	454	2 A10725	para-aminobenzoate
11	82	10.9	454	2 A10725	p-aminobenzoate sy
12	81.5	10.8	359	2 T35983	probable peptide t
13	81	10.7	195	2 T36975	hypothetical prote
14	80.5	10.7	366	2 D69506	probable 2-oxoisov
15	80	10.6	324	2 D63695	quinone oxidoreduc
16	80	10.6	764	2 H82717	malate oxidoreduct
17	79.5	10.5	110	2 G70922	hypothetical prote
18	79.5	10.5	148	2 B83291	conserved hypothet
19	79.5	10.5	657	2 S30288	C4-dicarboxylate s
20	79.5	10.5	2591	2 T30288	pristinamycin I sy
21	79	10.5	448	2 G70977	hypothetical prote
22	79	10.5	566	2 T35203	probable two-compo
23	79	10.5	998	2 T35745	probable ATP-bindi
24	78.5	10.4	919	2 T29581	hypothetical prote
25	78.5	10.4	1122	2 T42400	Eph receptor tyros
26	77.5	10.3	562	2 B70609	hypothetical prote
27	77	10.2	202	2 S18528	hypothetical prote
28	76.5	10.1	124	2 T36759	hypothetical prote
29	76.5	10.1	462	2 B75306	probable cycloprop

30	76.5	10.1	568	2 T28876	hypothetical prote
31	76.5	10.1	775	2 A32494	transposable eleme
32	76.5	10.1	3573	2 S23070	erythronolide synt
33	76	10.1	397	2 T35609	whiE protein I - S
34	75.5	10.0	254	2 JQ0490	regulatory protein
35	75.5	10.0	429	2 S72829	trehalose-6-phosph
36	75	9.9	286	2 B84226	3-hydroxyacyl-CoA
37	75	9.9	482	2 A38533	transcription acti
38	75	9.9	698	2 A82593	hypothetical prote
39	75	9.9	765	2 T49346	conserved hypothet
40	75	9.9	983	2 E86989	probable integral
41	75	9.9	992	2 G70950	hypothetical prote
42	74.5	9.9	195	2 C95286	probable transposa
43	74.5	9.9	370	2 C95374	probable ABC trans
44	74	9.8	368	2 G83015	3-dehydroquinatase
45	74	9.8	570	2 T30527	hypothetical prote

ALIGNMENTS

RESULT 1

A70850

hypothetical protein Rv0080 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: A70850

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Beldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: A70850

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-152 <COL>

A:Cross-references: UNIPROT:O53625; GB:AL021428; GB:AL123456; NID:g3261514; PIDN:CAA16261

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: Rv0080

C:Superfamily: Streptomyces coelicolor hypothetical protein SCJ12.09c

Query Match	Best Local Similarity	Score	DB 2;	Length
Matches 152;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;	

QY 1 MSPGRRASPOQAREVVELDRDEAMRLLASVDHGRVVFTRAAPAIRPVNHLVVDGRVIG 60

DB 1 MSPGRRASPOQAREVVELDRDEAMRLLASVDHGRVVFTRAAPAIRPVNHLVVDGRVIG 60

QY 61 RRLTAKVSVVRSSADAGVVVAYEADDLDPRRTGWSVVVTGLATEVSDPEQVARYQL 120

DB 61 RRLTAKVSVVRSSADAGVVVAYEADDLDPRRTGWSVVVTGLATEVSDPEQVARYQL 120

QY 121 LHPWNMAMDTVVAIEPIVTGIRIVADSRTP 152

DB 121 LHPWNMAMDTVVAIEPIVTGIRIVADSRTP 152

RESULT 2

T37028

hypothetical protein SCJ12.09c - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

R;Murphy, L.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, August 1999

A:Reference number: Z21619

A:Accession: T37028

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A;Residues: 1-146 <MUR>
A;Cross-references: UNIPROT:Q9R149; EMBL:AL109989; PIDN:CAB53421.1; GSPDB:GN00070; SCOPEDB:SCU12.09c
A;Experimental source: strain A3(2)
C;Generics:
A;Gene: SCOEDB:SCU12.09c
C;Superfamily: Streptococcus coelicolor hypothetical protein SCU12.09c

Query Match 29.9%; Score 225.5; DB 2; Length 146;
Best Local Similarity 42.9%; Pred. No. 7.1e-14;
Matches 57; Conservative 25; Mismatches 40; Indels 11; Gaps 4

QY 18 ELDRDEAMRLLASVDHGRVVFTRAALPAIRPVNH-LVVDGRVIGRTRLTAKVSVAVRSSA 76

Db 9 ELNRQESLQRLANAPVGR1VHTRDRLPAVLPNFVLEKSGAVLRTSSSELVRAV----- 64

QY 77 DAGVVAYEADDLDPRRRTGWSVVVTGLATEVSDP---EQVARYORLLLHPWVNMAMDTVV 133

Db 65 - DGAVVAFEADVDVATHSGSVVVTGLASVVTPDGHEQLVRTGP -- RSWVPWPVEVVF 121

QY 134 AIEPEIVTGIRIV 146

—
..
—
..
—
..
—
..
—

RESULT 3
T37043
hypothetical protein SCJ12.26 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T37043
R:Murphy, L.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A:Reference number: Z21619
A:Accession: T37043

Query Match 29.6%; Score 223.5; DB 2; Length 137;
Best Local Similarity 43.7%; Pred. No. 1e-13;
Matches 59; Conservative 13; Mismatches 44; Indels 19; Gaps 3

QY 19 LDRDEAMRLLASVDHGRVVFTRAALPAIRPVNHLVVDGRVIGRTRLTAKVSVAVRSSADA 78

```

Db      10 LERAECURLLGSAPVGRIVYTREALPAVLPVNF-----SLDTASVURTSGS 58

```

QY 79 -----GVVAYEADDLP RRRTGWSVVVTGLATEVSDPEQVARY-QRLLHPWVNMMAMD 130

Db 59 DLVRAVDGAVAEADAFDAHDRSGSVVVTGRATVVTDPAAKRLAENGPRSVWVSRDG 118

Qv 131 TWVAIEPEIVTGIRI 145

RESULT 4
T36956
probable DNA-binding protein - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T36956
R/Seeger, K.J.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1999
A/Reference number: Z21607
A/Accession: T36956
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-323 <SEE>

A;Cross-references: UNIPROT:Q9RIZ1; EMBL:AL109962; PTDN:CAB53141.1; GSPDB:GN00070; SCODED:
A;Experimental source: strain A3(2)
C;Generics:
A;Gene: SCOEDB:SCJ1.23C

Query Match 23.6%; Score 178; DB 2; Length 323;
Best Local Similarity 34.3%; Pred. No. 4.7e-09;
Matches 46; Conservative 19; Mismatches 61; Indels 8; Gaps 2;

18 ELDRDEAMRLLASVDHGRWTFTRAALPAIRPVNHLVWDGRVIGRTRLTAKVSAVRSSAD 77

Db 194 DLDEKECRALLSAHGVRRLAINTPDGFAVFPPLNYVVSOGTIAERTDPDPAATAAAVDSE-- 251

78 AGVVVAYEADDLPPRRRTGWSVVVT'GLATEVSDPEOVARYORLLH--PWVNMAMDT'VVAI 135

252 ----VAFVDHIDDMSQSVLAVGRCRAVTDPEEVRRLAEKAHTEPNAGGRRTMWLAV 307

Ov 136 EPEIVTGIRIVADS 149

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

RESULT 5
T36942
probable DNA-binding protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36942
R:Seeger, K.J.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z21607
A:Accession: T36942
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-233 <SEE>
A:Cross-references: UNIPROT:Q9RJ05; EMBL:AL109962; PIDN:CAB53127.1; GSFDB:GN00070; SCOEDE
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDE:SCJ1.09C

Query Match 17.4%; Score 131; DB 2; Length 233;
Best Local Similarity 32.3%; Pred. NO. 7.5e-05;
Matches 50; Conservative 19; Mismatches 68; Indels 18; Gaps 8;

QY 3 PGSRRASQSAREVVELDRDEAMRLLASVDHGRVFTRA--ALPAIRPVNHLVVDGR-V 58

Db 88 PG--GGGAAHPVLVRLSEQCWQRLCTHGIGRISYVAGPGKEAPVVPVNF--VDGRSV 144

QY 59 IGRTRLTAKVSAVRSSADAGVVAYEADLDPPRRRTGWSVVVTGLA--TEVSDPEQVA 115

Db 145 VYRTDPAGVAGIR-----ACEPVATEADHYVDENTGLGWSVLLAGTAEHPVEREALEALA 198

116 RYORLLHPVNNMAMDTVVAIEPEIVTGIRIVADSR 150 QY

199 R-RRGAVPWAGGRRDLWVRVLP HQVSG-RVIOPLR 231

RESULT 6
T36500 hypothetical protein SCGP3.20 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36500
R:Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1999
A:Reference number: Z21608
A:Accession: T36500
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-135 <SAU>
A:Cross-references: UNIPROT_Q9XA58; EMBL:AL096822; PIDN:CABA6942.1; GSPDB:GN000070; SCODE
A:Experimental source: strain A3 (2)
C:Genetics:

para-aminobenzoate synthase component I [imported] - Salmonella enterica subsp
 AI0725
 C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
 A:Note: this species has also been called *Salmonella typhi*
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AI0725
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.;
 th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.

S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: A10725
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-454 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD05507.1; PID:gl6503011; GSPDB:GN00176
C:Genetics:
A:Gene: STY1954
C:Superfamily: anthranilate synthase component I

Query Match 10.9%; Score 82; DB 2; Length 454;
Best Local Similarity 25.7%; Pred. No. 5.9;
Matches 44; Conservative 22; Mismatches 53; Indels 52; Gaps 10;

QY 7 RASPOSAREVVDELDRDEAMRLLASVDHGRVVFTRAAALP-----AIRP---VNHLV- 53
Db 285 QADROQAOKLANSMKDRLENLMI-VDLNRNDIGRVAVPGSVKVPFLFVVEFPFPAVHLLVS 343
QY 54 -VDGRVIGRTRLTAKVSVAVRSSADAGV--VVAYE-ADDDLPRRRTGW----- 97
Db 344 TITARLPDSLHATDLLRAAFPGSGITGAPKVRAMEIIDLEPQRNNAWCGSIGVLSFCGK 403
QY 98 ---SVVV-TGLATE-----VSDPEQVARYQ-----RLHHPWN 126
Db 404 MDTSTIRTVTATQOLYCSAGGGIVADSNBEEAYQETFDKVNRIHLPLEN 454

RESULT 11
A1132
p-aminobenzoate synthase (EC 4.1.3.-) component I - *Salmonella typhimurium*
C:Species: *Salmonella typhimurium*
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
A:Accession: A11132
R.Goncharoff, P.; Nichols, B. P.
Mol. Biol. Evol. 5, 531-548, 1988
A:Title: Evolution of aminobenzoate synthases: nucleotide sequences of *Salmonella typhim*
A:Reference number: A93063; MUID:89056707; PMID:3057324
A:Accession: A11132
A:Molecule type: DNA
A:Residues: 1-454 <GN>
A:Cross-references: UNIPROT:P12680; GB:M22079; NID:gl54230; PIDN:AAA88618.1; PID:gl54231
C:Genetics:
A:Gene: pabb
C:Superfamily: anthranilate synthase component I
C:Keywords: carbon-carbon lyase; oxo-acid-lyase

Query Match 10.9%; Score 82; DB 2; Length 454;
Best Local Similarity 25.7%; Pred. No. 5.9;
Matches 44; Conservative 22; Mismatches 53; Indels 52; Gaps 10;

QY 7 RASPOSAREVVDELDRDEAMRLLASVDHGRVVFTRAAALP-----AIRP---VNHLV- 53
Db 285 QADROQAOKLANSMKDRLENLMI-VDLNRNDIGRVAVPGSVKVPFLFVVEFPFPAVHLLVS 343
QY 54 -VDGRVIGRTRLTAKVSVAVRSSADAGV--VVAYE-ADDDLPRRRTGW----- 97
Db 344 TITARLPDSLHATDLLRAAFPGSGITGAPKVRAMEIIDLEPQRNNAWCGSIGVLSFCGK 403
QY 98 ---SVVV-TGLATE-----VSDPEQVARYQ-----RLHHPWN 126
Db 404 MDTSTIRTVTATQOLYCSAGGGIVADSNBEEAYQETFDKVNRIHLPLEN 454

RESULT 12
T35983
probable peptide transport ATP-binding protein - *Streptomyces coelicolor*
C:Species: *Streptomyces coelicolor*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T35983

R.Murphy, L.; Harris, D.; Bontley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, February 1999
A:Reference number: 221589
A:Accession: T35983
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-359 <MUR>
A:Cross-references: UNIPROT:Q92531; EMBL:AL035559; PIDN:CAB37471.1; GSPDB:GN00070; SCOEDI
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC92.04
C:Superfamily: inner membrane protein malk; ATP-binding cassette homology

Query Match 10.8%; Score 81.5; DB 2; Length 359;
Best Local Similarity 29.5%; Pred. No. 5;
Matches 31; Conservative 18; Mismatches 47; Indels 9; Gaps 5;

QY 3 PGRRRASPOSAREVVVELD--RDEAMRLLASVDHGRVVFTRAAALPAIRPV-NHLVVVDGRVI 59
Db 43 PGRGAAARAVDGVLDIRGEIVAVLGESGCKTTLARSLGLVVRTGGRVTFDGAFL 102
QY 60 GRTRLTAKVSVAVRSSADAGVVVAYEADDDLPRRRTGWSVVVTGL 104
Db 103 G---YSARALKAYRRRAQ--LVLQDPGSLNP-RHTVYDIVAEGL 141

RESULT 13
T36975
hypothetical protein SCJ11.04 - *Streptomyces coelicolor*
C:Species: *Streptomyces coelicolor*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36975
R.Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A:Reference number: 221618
A:Accession: T36975
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-195 <OLI>
A:Cross-references: UNIPROT:Q9BIA0; EMBL:AL109949; PIDN:CAB52889.1; GSPDB:GN00070; SCOEDI
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCJ11.04

Query Match 10.7%; Score 81; DB 2; Length 195;
Best Local Similarity 24.3%; Pred. No. 2.7;
Matches 34; Conservative 17; Mismatches 33; Indels 56; Gaps 5;

QY 4 GSRRASPOSAREVVVELDRDEAMRLLASVDHGRVVFTRAAALPAIRPVNHLVVVDGRVIGRTR 63
Db 22 GLLEAREASAREVEVLREEAARAVALAEAGEI-----ELDRRVIAREE 65
QY 64 LTAKVSV-----AVRSSADAGVVVAYEADDDLPRRRTGWSVVVT 102
Db 66 LVEALAVSAEATTGVTAEAGEGEGTALVPAPASAAEPGAIVPH-----WQ---E 110
QY 103 GLATEVSDPEQVARYQRLH 122
Db 111 GLSVSVLSPNN---QRILN 126

RESULT 14
D69506
probable 2-oxoisovalerate-ferredoxin oxidoreductase (EC 1.2.7.-) alpha chain AF2053 [nim]
C:Species: *Archaeoglobus fulgidus*
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: D69506
R.Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,
F.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.N.
Smith, H.O.; Woese, C.R.; Venter, J.C.

Search completed: October 15, 2005, 04:38:19
Job time : 9.48147 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:24:25 ; Search time 34.047 Seconds
(without alignments)
2286.133 Million cell updates/sec

Title: US-10-617-038-2

Perfect score: 755

Sequence: 1 MSPGRRASQSAREVVELD.....VAIEPIVTGIRIVADSRTP 152

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Uniprot_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	755	100.0	152	2	O53625 mycobacteri
2	746	98.8	152	2	Q7U2W2 mycobacteri
3	678	89.8	137	2	Q7DAH5 mycobacteri
4	225.5	29.9	146	2	Q9RI49 streptomyc
5	223.5	29.6	137	2	Q9RI34 streptomyc
6	205	27.2	208	2	Q82NN1 streptomyc
7	178	23.6	323	2	Q9RI21 streptomyc
8	144	19.1	151	2	Q73V39 mycobacteri
9	131	17.4	233	2	Q9RI05 streptomyc
10	130	17.2	140	2	Q6A7Q2 propionibac
11	127	16.8	138	2	Q73U04 mycobacteri
12	121	16.0	137	2	Q6M2U2 corynebacte
13	119	15.8	163	2	Q8FMS0 corynebacte
14	103	13.6	135	2	Q9XA58 streptomyc
15	100	13.2	145	2	Q6NFL3 corynebacte
16	98.5	13.0	164	2	Q82F88 streptomyc
17	96	12.7	159	2	Q6ACG4 leifsonia x
18	89	11.8	237	2	Q7NNI2 mycobacteri
19	88.5	11.7	437	1	UDG_RHIME
20	88	11.7	186	2	Q98DF8 rhizobium m
21	86	11.4	945	2	Q92Q30 rhizobium l
22	85.5	11.3	650	1	GSPD_ECOLI
23	85	11.3	307	2	Q59Z73 streptomyc
24	84.5	11.2	237	2	Q73ZG1 mycobacteri
25	84.5	11.2	11096	2	Q9LIW3 streptomyc
26	84	11.1	380	1	SCHA_STRHA
27	84	11.1	577	2	Q6N351 rhodopseudo
28	84	11.1	764	2	Q87EN1 xylella fas
29	83.5	11.1	251	1	TPIS_TRYCR
30	83.5	11.1	419	2	Q9UIB9 leishmania
31	83.5	11.1	654	2	Q8CVN5 escherichia

32	83	11.0	260	2	Q9FBT1 streptomyc
33	83	11.0	480	2	Q7NT82 chromobacte
34	83	11.0	778	2	Q9FBP3 streptomyc
35	82.5	10.9	442	2	Q6W4V7 rhizobium l
36	82.5	10.9	852	2	Q7Y097 oryza sativ
37	82	10.9	391	2	Q6D1D6 erwinia car
38	82	10.9	454	1	PABB_SALTY
39	82	10.9	454	2	Q8Z677 salmonella
40	81.5	10.8	302	2	Q7NI68 glosobacter
41	81.5	10.8	359	2	Q9Z531 streptomyc
42	81.5	10.8	577	2	Q62C53 burkholderi
43	81	10.7	195	2	Q9RIAO streptomyc
44	81	10.7	527	2	Q6C8A9 yarrowia li
45	80.5	10.7	329	2	Q8N225 homo sapien

ALIGNMENTS

RESULT 1
O53625 PRELIMINARY; PRT; 152 AA.
AC O53625;
DT 01-JUN-1998 (TREMELrel. 06, Created)
DT 01-JUN-1998 (TREMELrel. 06, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=Rv0080;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H378V;
RX MEDLINE=9895987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
RL Nature 393:537-544(1998).
DR EMBL; BX842572; CAA18261.1; -.
DR PIR; A70850; A70850.
DR Tuberculist; RV0080; -.
DR InterPro; IPR009002; FMN binding.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 152 AA; 16574 MW; 6EAD801AE5F21C67 CRC64;

Query Match 100.0%; Score 755; DB 2; Length 152;
Best Local Similarity 100.0%; Pred. No. 6.9e-60;
Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSPGRRASQSAREVVELDRDREANRLASVDHGRVVFTRAAALPAIRPVNHLVVDGRVIC 60
DB 1 MSPGRRASQSAREVVELDRDREANRLASVDHGRVVFTRAAALPAIRPVNHLVVDGRVIC 60
QY 61 RRLTAKVAVVRSSADAGVWVAYEADDLPDRRTGWSVVVTGLATEVSDPEQVARYQRL 120
DB 61 RRLTAKVAVVRSSADAGVWVAYEADDLPDRRTGWSVVVTGLATEVSDPEQVARYQRL 120
QY 121 LHPWNVMAMDTTVAIEPIVTGIRIVADSRTP 152
DB 121 LHPWNVMAMDTTVAIEPIVTGIRIVADSRTP 152

RESULT 2
Q7U2W2

```

ID Q7U2W2 PRELIMINARY; PRT; 152 AA.
AC Q7U2W2;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein Mb0083.
GN OrderedLocusNames=Mb0083;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garner T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248334; CAD92945.1; -.
DR InterPro; IPR009002; FMN binding.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 152 AA; 16616 MW; 6F0681AASF30C76 CRC64;

Query Match 98.8%; Score 746; DB 2; Length 152;
Best Local Similarity 99.3%; Pred. No. 4.4e-59;
Matches 151; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSPGRRASPSQAREVELDRDEAMRLASVDHGRVWFTFRAALPAIRPVNHLVVDGRVIG 60
DB 1 MSPGRRASPSQAREVELDRDEAMRLASVDHGRVWFTFRAALPAIRPVNHLVVDGRVIV 60

QY 61 RTRLTAKVSVAVRSADAGVWVAYEADDLPDRRTGWSVVVTGLATEVSDPEQVARYQRL 120
DB 61 RTRLTAKVSVAVRSADAGVWVAYEADDLPDRRTGWSVVVTGLATEVSDPEQVARYQRL 120

QY 121 LHPWNMAMDTVVAIEPEIVTGIRIVADSRTP 152
DB 121 LHPWNMAMDTVVAIEPEIVTGIRIVADSRTP 152

RESULT 3
Q7DAH5 PRELIMINARY; PRT; 137 AA.
ID Q7DAH5
AC Q7DAH5;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MT0087;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Uitterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
DR EMBL; AE000516; AAK44312.1; -.
DR TIGR; MT0087; -.
DR InterPro; IPR009002; FMN_binding.

```

```

KW Hypothetical protein.
SQ SEQUENCE 137 AA; 15007 MW; 9BE0834105D07C2B CRC64;

Query Match 89.8%; Score 678; DB 2; Length 137;
Best Local Similarity 99.3%; Pred. No. 4.9e-53;
Matches 136; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 VVELDRDEAMRLASVDHGRVWFTFRAALPAIRPVNHLVVDGRVIGRTRLTAKVSVAVRS 75
DB 1 VVELDRDEAMRLASVDHGRVWFTFRAALPAIRPVNHLVVDGRVIGRTRLTAKVSVAVRS 60

QY 76 ADAGVWVAYEADDLPDRRTGWSVVVTGLATEVSDPEQVARYQRLHLPWNMAMDTVVAI 135
DB 61 ADAGVWVAYEADDLPDRRTGWSVVVTGLATEVSDPEQVARYQRLHLPWNMAMDTVVAI 120

QY 136 EPEIVTGIRIVADSRTP 152
DB 121 EPEIVTGIRIVADSRTP 137

RESULT 4
Q9RI49 PRELIMINARY; PRT; 146 AA.
ID Q9RI49
AC Q9RI49;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein SC00197.
GN ORFNames=SCJ12.09c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(12) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939104; CAB53421.1; -.
DR PIR; T37028; T37028.
DR InterPro; IPR009002; FMN binding.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 146 AA; 15808 MW; CC01277D453BD5EC CRC64;

Query Match 29.9%; Score 225.5; DB 2; Length 146;
Best Local Similarity 42.9%; Pred. No. 2e-12;
Matches 57; Conservative 25; Mismatches 40; Indels 11; Gaps 4;

QY 18 ELDRDEAMRLASVDHGRVWFTFRAALPAIRPVNH-LVVDGRVIGRTRLTAKVSVAVRS 76
DB 9 ELNRQESLQRLANAPVGRIVHTRDALPAVLPVNFVLEKSGAVLLRTSASSELVRAV---- 64

QY 77 DAGVVVAYEADDLPDRRTGWSVVVTGLATEVSDP---EQVARYQRLHLPWNMAMDTV 133
DB 65 -DGAVVAFEADVDVATHSGWSVVVTGLASVVVTDFGEHEQLVRTGP--RSWVPNPVEVPV 121

QY 134 AIEPEIVTGIRIV 146
DB 122 RIAPDLVSGRELIV 134

RESULT 5

```

metabolites.";
Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).
[2]
SEQUENCE FROM N.A.
STRAIN=MA-4680;
MEDLINE=22608306; PubMed=12692562;
Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
Sakaki Y., Hattori M., Omura S.;
RA "Complete genome sequence and comparative analysis of the industrial
microorganism Streptomyces avermitilis.";
Nat. Biotechnol. 21:526-531(2003).
RL EMBL; AP005026; BAC68981.1; --
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR009002; FMN binding.
DR InterPro; IPR001387; HTH 3.
DR InterPro; IPR010982; Lambda like DNA.
DR InterPro; IPR011576; Pyridox_ox_like.
DR Pfam; PF01381; HTH 3, 1.
DR Pfam; PF01243; Pyridox_oxidase; 1.
DR SMART; SM00530; HTH_XRE; 1.
DR PROSITE; PS00943; HTH_CROCI; 1.
KW Complete proteome; DNA-binding.
SQ SEQUENCE 208 AA; 22236 MW; AF06C031D574E3D5 CRC64;

Query Match 27.2%; Score 205; DB 2; Length 208;
Best Local Similarity 37.2%; Pred. No. 2.1e-10;
Matches 55; Conservative 22; Mismatches 61; Indels 10; Gaps 3;

QY 1 MSPGSRASPOSAREVVELDRDEAMRLASVDHGRVVFTRALPAIRPVNHLVVDGRVIG 60
Db 67 LPPGVGRAAHNP--ELLESDDEECARLSTHGVRGLAVDTPTGPIVPLNYSVVDGVVCF 124
QY 61 RTRLTAKVSVAVRSGSADAGVVVAYEADDLPDRRTGMSVVVTGLATEVSDPEQVARYQL 120
Db 125 RTAADSE-----PAASAGSRVAFVVDHIDEALSQGSLVRLGRLVDTPTDTRLRLAEL 178
QY 121 LH--PWNVAMDTVVAIPEIVTGIRIV 146
Db 179 AYSGPWAGGERDTWVCDPVGVTGRRIV 206

RESULT 7
Q9RI31 PRELIMINARY; PRT; 323 AA.
AC Q9RI31; (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Putative DNA-binding protein.
GN ORFNames=SCJ1.23G;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
Huang C.-H., Kieser T., Latke L., Murphy L.D., Oliver K., O'Neill S.,
Rabinowitch E., Rajandream M.A., Rutherford K.M., Rutter S.,
Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
Hopwood D.A.;
RA "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
Nature 417:141-147(2002).
RL EMBL; AL939104; CAB53438.1; --
DR PIR; T37043; T37043.
DR InterPro; IPR009002; FMN binding.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 137 AA; 14532 MW; 0D1E5EC4D84C78A8 CRC64;

Query Match 29.6%; Score 223.5; DB 2; Length 137;
Best Local Similarity 43.7%; Pred. No. 2.8e-12;
Matches 59; Conservative 13; Mismatches 44; Indels 19; Gaps 3;

QY 19 LDRDEAMRLASVDHGRVVFTRALPAIRPVNHLVVDGRVIGRTRLTAKVSVAVRSGSADA 78
Db 10 LERAELRLLSAPGRIVYTRALPVLVNF-----SLQDASVLRISAGS 58
QY 79 -----GVVAYEADDLPDRRTGMSVVVTGLATEVSDPEQVARY-QLLHPWVNNAMD 130
Db 59 DLVRAVDGAVAFADAFDHRSGMSVVVTGTRATVTDPAARLARLAENGPRSNVWSRDG 118
QY 131 TVVAIPEIVTGIRI 145
Db 119 LVYRIESEMVTGREI 133

RESULT 6
Q82NN1 PRELIMINARY; PRT; 208 AA.
AC Q82NN1; (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Putative DNA-binding protein.
GN OrderedLocusNames=SAV1271;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RA "Genome sequence of an industrial microorganism Streptomyces
avermitilis: deducing the ability of producing secondary
metabolites.";
Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).
[2]
SEQUENCE FROM N.A.
STRAIN=MA-4680;
MEDLINE=22608306; PubMed=12692562;
Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
Sakaki Y., Hattori M., Omura S.;
RA "Complete genome sequence and comparative analysis of the industrial
microorganism Streptomyces avermitilis.";
Nat. Biotechnol. 21:526-531(2003).
RL EMBL; AP005026; BAC68981.1; --
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR009002; FMN binding.
DR InterPro; IPR001387; HTH 3.
DR InterPro; IPR010982; Lambda like DNA.
DR InterPro; IPR011576; Pyridox_ox_like.
DR Pfam; PF01381; HTH 3, 1.
DR Pfam; PF01243; Pyridox_oxidase; 1.
DR SMART; SM00530; HTH_XRE; 1.
DR PROSITE; PS00943; HTH_CROCI; 1.
KW Complete proteome; DNA-binding.
SQ SEQUENCE 208 AA; 22236 MW; AF06C031D574E3D5 CRC64;

Query Match 27.2%; Score 205; DB 2; Length 208;
Best Local Similarity 37.2%; Pred. No. 2.1e-10;
Matches 55; Conservative 22; Mismatches 61; Indels 10; Gaps 3;

QY 1 MSPGSRASPOSAREVVELDRDEAMRLASVDHGRVVFTRALPAIRPVNHLVVDGRVIG 60
Db 67 LPPGVGRAAHNP--ELLESDDEECARLSTHGVRGLAVDTPTGPIVPLNYSVVDGVVCF 124
QY 61 RTRLTAKVSVAVRSGSADAGVVVAYEADDLPDRRTGMSVVVTGLATEVSDPEQVARYQL 120
Db 125 RTAADSE-----PAASAGSRVAFVVDHIDEALSQGSLVRLGRLVDTPTDTRLRLAEL 178
QY 121 LH--PWNVAMDTVVAIPEIVTGIRIV 146
Db 179 AYSGPWAGGERDTWVCDPVGVTGRRIV 206

RESULT 7
Q9RI31 PRELIMINARY; PRT; 323 AA.
AC Q9RI31; (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Putative DNA-binding protein.
GN ORFNames=SCJ1.23G;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
Huang C.-H., Kieser T., Latke L., Murphy L.D., Oliver K., O'Neill S.,
Rabinowitch E., Rajandream M.A., Rutherford K.M., Rutter S.,
Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
Hopwood D.A.;
RA "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
Nature 417:141-147(2002).
RL EMBL; AL939104; CAB53438.1; --
DR PIR; T37043; T37043.
DR InterPro; IPR009002; FMN binding.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 137 AA; 14532 MW; 0D1E5EC4D84C78A8 CRC64;

```

DR InterPro; IPR009002; FMN binding.
DR InterPro; IPR001387; HTH_3.
DR InterPro; IPR010982; Lambda_like_DNA.
DR InterPro; IPR011576; Pyridox_ox_like.
DR Pfam; PF01381; HTH_3; 1.
DR Pfam; PF01243; Pyridox_oxidase; 1.
DR SMART; SM00530; HTH_XRE; 1.
DR PROSITE; PS0943; HTH_CROCI; 1.
KW Complete proteome; DNA-binding.
SQ SEQUENCE 323 AA; 34180 MW; 2F1530BAC848E67 CRC64;

Query Match 23.6%; Score 178; DB 2; Length 323;
Best Local Similarity 32.3%; Pred. No. 9e-08;
Matches 46; Conservative 19; Mismatches 61; Indels 8; Gaps 2;

QY 18 ELDRDEAMRLASVDHGRVVFTRAAALPAIRPNVHVLVDGRVIGRTRLTAKVSVAVRRSAD 77
DB 194 DLDEKCRALLSAHGVGRALNTPDPAVFLNYSQGTIAFRTPDPAATAAAVDS- 251
QY 78 AGVVVAYEADDLPDRRTGWSVVVTGLATEVSDPEQVARYQRLH--PWNVMMMDTVVAI 135
DB 252 ----VAFEVDHIDDPMSQGSVLAVGRCRAVTPDPEVRRLAEKAHTEPWAGGRRTMWLAV 307

QY 136 EPEIVTGIRIVADS 149
DB 308 ETERLTGRTITSVS 321

RESULT 8
Q73V39
ID Q73V39 PRELIMINARY; PRT; 151 AA.
AC Q73V39
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=Map3177;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Anonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017239; AAS05725.1; -.
DR InterPro; IPR009002; FMN binding.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 151 AA; 17022 MW; 16B69E11D2EC097F CRC64;

Query Match 19.1%; Score 144; DB 2; Length 151;
Best Local Similarity 37.3%; Pred. No. 4.2e-05;
Matches 50; Conservative 10; Mismatches 66; Indels 8; Gaps 3;

QY 12 SAREVVELDRDEAMRLASVDHGRVVFTRAAALPAIRPNVHVLVDGRVIGRTRLTAKUSVA 71
DB 11 SDEPVQVISEBNWRLGSLVALGRVLTWTFAGEAEIFFNVVQDRTVLFRTAEGTKLFSA 70
QY 72 VRSSADAGVVVAYEADDLPDRRTGWSVVVTGLATEVSDPEQVARYQRL-LHPVNMAMD 130
DB 71 VANHA-----VVFVFD--DHNVLVEGMSVIVGRARLLKTDTDIQRAERALLPWTATLKP 123
QY 131 TVVAIEPEIVTGIR 144
DB 124 HYVITPTVTGRR 137

RESULT 9
Q9RJ05
ID Q9RJ05 PRELIMINARY; PRT; 233 AA.
AC Q9RJ05
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

```

```

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
GN Possible DNA-binding protein.
GN ORFNames=SCJ1.03c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornaby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrall B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RL Nature 417:141-147(2002).
DR EMBL; AL919104; CAB53127.1; -.
DR PIR; T36942; T36942.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR009002; FMN binding.
DR InterPro; IPR001387; HTH_3.
DR InterPro; IPR010982; Lambda_like_DNA.
DR Pfam; PF01381; HTH_3; 1.
DR SMART; SM00530; HTH_XRE; 1.
DR PROSITE; PS0943; HTH_CROCI; 1.
KW Complete proteome; DNA-binding.
SQ SEQUENCE 233 AA; 25319 MW; 6AAF3A97971E9DE4 CRC64;

Query Match 17.4%; Score 131; DB 2; Length 233;
Best Local Similarity 32.3%; Pred. No. 0.001;
Matches 50; Conservative 19; Mismatches 68; Indels 18; Gaps 8;

QY 3 PGSRRASPOSAREVVELDRDEAMRLASVDHGRVVFTRAAA--ALPAIRPNVHVLVDGR-V 58
DB 88 PG--RGGPAAHPVLRLSEQECWQRLCTHGRTSYVAGPKEAPVVPVNF--VDGRSV 144
QY 59 IGRTRLTAKVSVAVRRSADAGVVVAYEADDLPDRRTGWSVVVTGLA---TEVSDPEQVA 115
DB 145 VYRTDPAGVAGIR-----AGEPVAFEDHVDGMTGLGWSVLLAGTAEHPVPERALEALA 198
QY 116 RYQRLHHPVNMAMDTVVAIEPEIVTGIRIVADSR 150
DB 199 R-RGAVPWAGGRRLDWVRVLPVQVSG-RVIQPLR 231

RESULT 10
Q6A702
ID Q6A702 PRELIMINARY; PRT; 140 AA.
AC Q6A702;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Conserved protein.
GN OrderedLocusNames=PPA1465;
OS Propionibacterium acnes.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Propionibacterineae; Propionibacteriaceae; Propionibacterium.
OX NCBI_TaxID=1747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KPA171202 / DSM 16379;
RX PubMed=15286373; DOI=10.1126/science.1100330;
RA Brueggemann H., Henne A., Hofer F., Liesegang H., Wieroz A.,
RA Strittmatter A., Hujer S., Duerre P., Gottschalk G.;
RT "The complete genome sequence of Propionibacterium acnes, a commensal

```

[illegible]

```

Query Match      15.8%; Score 119; DB 2; Length 163;
Best Local Similarity 32.4%; Pred. No. 0.0081;
Matches 46; Conservative 15; Mismatches 73; Indels 8; Gaps 4;

QY 2 SPGRSRASPOSAREVVELDRDDEAMRLASVDHGRVVFTRAAALPAIRPVNHLVVDGRVIGR 61
DB 17 SPGSVAASDVNGPQELTEASRERLARTTLGRVVVRGGDDLDLPFNYVTNGENLYFR 76

QY 62 TRLTAKVSVAVRSSADAGVVVAYEADDDLPDRRTGWSVVVTGLATEVSDPEQVARYQRL- 120
DB 77 TAEGNKL-FTITLND----VLFEDEV--RDGVAWSVVVRGDABLVQDFKEIQEVDEL 129

QY 121 LHPWNMAMDTVTVAPEIVTG 142
DB 130 LKPWPTLKYNVTRTPNEITG 151

RESULT 14
Q9XA58 PRELIMINARY; PRT; 135 AA.
AC Q9XA58;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein SC03819.
GN ORFNames=SCG3.20;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidaigo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939117; CAB46942.1; -.
DR PIR; T36500; T36500.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 135 AA; 14598 MW; CASF00B10188165B CRC64;

Query Match      13.6%; Score 103; DB 2; Length 135;
Best Local Similarity 33.7%; Pred. No. 0.18;
Matches 34; Conservative 10; Mismatches 35; Indels 22; Gaps 4;

QY 20 DRDEAMRLASVDHGRVVFTRAAALPAIRPVNHLVVDGRVI-----GRTRLTAKVSVAV 72
DB 4 DEQLAVGLLGRYAGRAATTLRALPFLAHARHIVADGRVLLRMPRSWGVRVC----- 56

QY 73 RSSADAGVVVAYEADDDLPDR--RTGWSVVVTGLATEVSDP 111
DB 57 -----AGSVVAYGADNLISARSGESLWTVQVVG-RCETHEP 91

RESULT 15
Q6NFL3 PRELIMINARY; PRT; 145 AA.
AC Q6NFL3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=DIP1874;

```

```

OS Corynebacterium diphtheriae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1717;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Biotype Gravis / NCTC 13129;
RX MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/gkg874;
RA Cerdano-Tarraga A.-M., Efstathiou A., Dover L.G., Holden M.T.G.,
RA Pallen M.J., Bentley S.D., Besra G.S., Churcher C.M., James K.D.,
RA De Zeyea A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
RA Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,
RA Rabinowitch E., Rutherford K.M., Thomson N.R., Unwin L.,
RA Whitehead S., Barrell B.G., Parkhill J.;
RT "The complete genome sequence and analysis of Corynebacterium
RT diphtheriae NCTC13129.";
RL Nucleic Acids Res. 31:6516-6523(2003).
DR EMBL; BX248359; CAES0403.1; -.
DR InterPro; IPR009002; FMN binding.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 145 AA; 16354 MW; 6D8D8954D4FAP511 CRC64;

Query Match      13.2%; Score 100; DB 2; Length 145;
Best Local Similarity 33.1%; Pred. No. 0.36;
Matches 43; Conservative 16; Mismatches 57; Indels 14; Gaps 6;

QY 19 LDRDEAMRLASVDHGRVVFTRAAALPAIRPVNHLVVD-----GRVIGTRLTAKVSVAVR 73
DB 12 LDSSDSLRLSSESVGRLVVHRKDDLDIFPVN-FVLDYSABEQPRVYFRTAEGTKL-FSVN 69

QY 74 SSADAGVVVAYEADDDLPDRRTGWSVVVTGLATEVSDPEQVARYQRL-LHPWNMAMDTV 132
DB 70 LNSD-----VLFEVDRFDDAE--GMSVLKGNAYVVRDTEEARHADTLGLKPWLPTLKYNF 123

QY 133 VAIEPEIVTG 142
DB 124 VRIDREVSG 133

Search completed: October 15, 2005, 04:35:48
Job time : 35.047 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:20:35 ; Search time 30.3061 Seconds
(without alignments)
. 1454.845 Million cell updates/sec

Title: US-10-617-038-3

Perfect score: 539

Sequence: 1 VESEPLYKLKAEFFKTLAHP.....RVLSDRVAVLELRAGGSAT 114

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	539	100.0	114	8	ADI37282 M. tuberc
2	251.5	46.7	113	3	AAB07556 Protein e
3	194	36.0	104	4	AAG90718 C glutami
4	145.5	27.0	275	4	AAB86528 P. pantot
5	141.5	26.3	218	6	ABU36669 Protein e
6	135.5	25.1	226	6	ABU36443 Protein e
7	135.5	25.1	226	6	ABU34625 Protein e
8	131.5	24.4	106	6	ABM71883 Staphyloc
9	127.5	23.7	95	6	ABU28998 Protein e
10	124.5	23.1	107	6	ABU40771 Protein e
11	124.5	23.1	121	7	ADF06894 Bacterial
12	123.5	22.9	105	6	ABU49708 Protein e
13	123	22.8	96	6	ABU18228 Protein e
14	122	22.6	107	6	ADB08342 Allostoc
15	119.5	22.2	108	6	ABU49198 Protein e
16	119	22.1	113	4	AAG82142 S. epider
17	119	22.1	124	5	ABP39352 Staphyloc
18	119	22.1	124	8	ADS04517 Staphyloc
19	114.5	21.2	118	4	AAB96258 Putative
20	114.5	21.2	124	6	ABU21809 Protein e
21	114	21.2	101	5	ABBA9983 Listeria
22	113.5	21.1	118	6	ABP57443 Mycobacte
23	112.5	20.9	123	4	AAG92245 C glutami
24	112.5	20.9	123	7	ADD13998 C. glutam
25	112.5	20.9	140	6	ABU35833 Protein e

ALIGNMENTS

RESULT 1

ADI37282	AD	ADI37282 standard; protein; 114 AA.			
XX	XX				
AC	AC	ADI37282;			
XX	XX				
DT	DT	22-APR-2004 (first entry)			
XX	XX				
DE	DE	M. tuberculosis low oxygen induced antigen RV0081 SEQ ID NO:3.			
XX	XX				
KW	KW	mycobacterial infection; vaccine; tuberculosis;			
KW	KW	Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;			
XX	XX	low oxygen induced antigen.			
XX	XX				
OS	OS	Mycobacterium tuberculosis.			
XX	XX				
PN	PN	WO2004006952-A2.			
XX	XX				
PD	PD	22-JAN-2004.			
XX	XX				
PF	PF	08-JUL-2003; 2003WO-DK000477.			
XX	XX				
PR	PR	13-JUL-2002; 2002DK-00001098.			
XX	XX	(STAT-) STATENS SERUM INST.			
XX	XX	Andersen P, Rosenkrands I, Stryhn A;			
XX	XX				
DR	DR	WPI: 2004-122778/12.			
DR	DR	N-PSDB; ADI37327.			
XX	XX				
PT	PT	Use of one or more polypeptides or their fragments, which are expressed			
PT	PT	during the latent stage of the mycobacterial infection, and/or nucleic			
PT	PT	acids encoding the polypeptides, for a therapeutic vaccine against			
XX	XX	tuberculosis.			
PS	PS	Claim 3; SEQ ID NO 3; 76pp; English.			
XX	XX				
CC	CC	The present invention describes polypeptides or their fragments, which			
CC	CC	are expressed during the latent stage of a mycobacterial infection,			
CC	CC	and/or nucleic acids encoding the polypeptides, which are useful for			
CC	CC	creating a therapeutic vaccine against tuberculosis. Also described: (1)			
CC	CC	a therapeutic vaccine against tuberculosis comprising one or more			
CC	CC	polypeptides; (2) a method for treating an animal, including a human			
CC	CC	being, with tuberculosis caused by virulent mycobacteria, e.g. by			
CC	CC	Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for			
CC	CC	immunising an animal, including a human being, against tuberculosis			
CC	CC	caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis			

Au61884 Propionib
Abm58403 Propionib
Abu34202 Protein e
Abu17579 Protein e
Abu44504 Protein e
Abu22386 Protein e
Abu35260 Protein e
Abu19574 Protein e
Abu23543 Protein e
Abu23543 Protein e
Abu24751 Protein e
Abu15460 Protein e
Abu33783 Protein e
Abu24235 Protein e
Abb4952 Lactococc
Abu18519 Protein e
Abo71152 Pseudomon
Adn47394 Thermococ
Abu16784 Protein e
Ada36437 Acinetoba

26 111.5 20.7 144 4 AAU61884
27 111.5 20.7 144 6 ABM58403
28 110.5 20.5 134 6 ABU34202
29 105 19.5 85 6 ABU17579
30 104 19.3 92 6 ABU44504
31 104 19.3 115 6 ABU22386
32 104 19.3 121 6 ABU35260
33 102.5 19.0 107 6 ABU19574
34 101.5 18.8 116 6 ABU23543
35 101 18.7 112 6 ABU23543
36 100 18.6 93 6 ABU24751
37 100 18.6 99 6 ABU15460
38 99.5 18.5 109 6 ABU33783
39 98 18.2 122 6 ABU24235
40 96.5 17.9 99 5 ABB4952
41 96 17.8 91 6 ABU18519
42 96 17.8 348 7 ABO71152
43 95.5 17.7 131 8 ADN47394
44 94.5 17.5 102 6 ABU16784
45 93.5 17.3 103 6 ADA36437

CC caused by virulent mycobacteria in an animal, including a human being;
CC (5) a method for diagnosing previous or ongoing infection with a virulent
CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
CC infection in a subject. The polypeptides have antibacterial activities,
CC and can be used in vaccines and in gene therapy. The polypeptides are
CC useful for the manufacture of a therapeutic vaccine for treating an
CC individual who is infected by a virulent mycobacterium, e.g. M.
CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
CC The present sequence represents a low oxygen induced antigen, which is
CC used in the exemplification of the present invention.

XX SQ Sequence 114 AA;

Query Match 100.0%; Score 539; DB 8; Length 114;
Best Local Similarity 100.0%; Pred. No. 1.2e-56;
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VESEPLYLKAEFFKTLAHPARIRILELLVERDRSVGELSSDVGLESSLNLSQQLGVLR 60
DB 1 VESEPLYLKAEFFKTLAHPARIRILELLVERDRSVGELSSDVGLESSLNLSQQLGVLR 60
QY 61 AGVAARRDGNAMYISIAAPDIAELLAVARKVLARVLSDRVAVLEDLRAGGSAT 114
DB 61 AGVAARRDGNAMYISIAAPDIAELLAVARKVLARVLSDRVAVLEDLRAGGSAT 114

RESULT 2
AAB07556
ID AAB07556 standard; protein; 113 AA.
XX AC AAB07556;
XX DT 20-OCT-2000 (first entry)
XX DE Protein encoded by the bleomycin (BLM) gene cluster ORF30.
XX KW BLM gene cluster; bleomycin gene cluster; polyketide metabolite;
XX KW bleomycin; bleomycin analogue; holo-carrier protein; thiazolidine;
XX KW thiazoline; bithiazoline; microbial metabolite; sugar.
XX OS Streptomyces verticillus.
XX FH Key Location/Qualifiers
XX FT Misc-difference 1 /note= "encoded by GTG"
XX PN WO200040704-A1.
XX PD 13-JUL-2000.
XX PF 06-JAN-2000; 2000WO-US000445.
XX PR 06-JAN-1999; 99US-0115435P.
XX PR 05-FEB-1999; 99US-0118848P.
XX PR 05-JAN-2000; 2000US-00477962.
XX PA (REGC) UNIV CALIFORNIA.
XX PI Shen B, Du L, Sanchez C, Chen M, Edwards DJ;
XX PI WPI; 2000-465974/40.
XX DR N-PSDB; AAA58471.
XX PT New bleomycin gene cluster components useful for peptide and/or
XX PT polyketide metabolites, especially bleomycin, production and for
XX PT chemically modifying biological molecules.
XX PS Disclosure; Page 97-137; 162pp; English.
XX CC AAB07556-78 represent proteins encoded by open reading frames (ORFs) 8 to
XX CC 30 of the BLM (Bleomycin) gene cluster. The proteins encoded by the gene
XX CC cluster are useful for producing peptides and/or polyketide metabolites,
XX CC especially bleomycin or bleomycin analogues. They are also useful for

CC chemically modifying biological molecules to produce branched methyl
CC groups, and for coupling amino acids and fatty acids. They may be reacted
CC with an apo-carrier protein and coenzyme A to produce a holo-carrier
CC protein. The BLM gene cluster or catalytic domains can be used
CC individually or collectively to produce thiazolidine, thiazoline,
CC bithiazoline and bithiazoline-containing microbial metabolites. The BLM
CC gene cluster may also be used to produce sugars

XX SQ Sequence 113 AA;

Query Match 46.7%; Score 251.5; DB 3; Length 113;
Best Local Similarity 48.1%; Pred. No. 4.9e-22;
Matches 50; Conservative 26; Mismatches 27; Indels 1; Gaps 1;
QY 5 PLYKLKAEFFKTLAHPARIRILELLVERDRSVGELSSDVGLESSLNLSQQLGVLRAGVV 64
DB 4 PLYOKAKAEFFMLGHPVRIRVLELLQDGPMPVRDLAA-IEIPSAISQQLAVLRSGIV 62
QY 65 AARRDGNAMYISIAAPDIAELLAVARKVLARVLSDRVAVLEDLR 108
DB 63 TSTRGTGVVYELAGGVAELMSAARRILTEMLNGQHLEELR 106

RESULT 3
AAG90718
ID AAG90718 standard; protein; 104 AA.
XX AC AAG90718;
XX DT 26-SEP-2001 (first entry)
XX DE C glutamicum protein fragment SEQ ID NO: 4472.
XX KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
XX KW organic acid synthesis.
XX OS Corynebacterium glutamicum.
XX PN EP1108790-A2.
XX PD 20-JUN-2001.
XX PF 18-DEC-2000; 2000EP-00127688.
XX PR 16-DEC-1999; 99JP-00377484.
XX PR 07-APR-2000; 2000JP-00159162.
XX PR 03-AUG-2000; 2000JP-00280988.
XX PA (KYOW) KYOWA HAKKO KOGYO KK.
XX PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX DR WPI; 2001-376931/40.
XX DR N-PSDB; AAH65937.
XX PT Novel polynucleotides derived from Coryneform bacteria, for identifying
XX PT mutation point of a gene, measuring expression of a gene, analyzing
XX PT expression profile or pattern of a gene and identifying homologous gene.
XX PS Claim 17; SEQ ID NO 4472; 246pp + Sequence Listing; English.
XX CC The present invention provides a number of nucleotide and protein
XX CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
XX CC are useful for identifying the mutation point of a gene derived from a
XX CC mutant of coryneform bacterium, measuring expression amount and analysing
XX CC the expression profile or expression pattern of a gene derived from
XX CC Coryneform bacterium, and identifying a homologue of a gene derived from
XX CC coryneform bacterium. Coryneform bacteria are useful for producing amino
XX CC acids, nucleic acids, vitamins, saccharides and organic acids,
XX CC particularly L-lysine. The present sequence is a protein described in the
XX CC exemplification of the invention. Note: The sequence data for this patent
XX CC did not form part of the printed specification, but was obtained in

the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

[illegible]

RESULT 8	
ABM71883	
ID	ABM71883 standard; protein; 106 AA.
XX	
XX	
AC	ABM71883;
XX	
DT	20-NOV-2003 (first entry)
XX	
XX	
DE	Staphylococcus aureus protein #1123.
XX	
XX	
KW	Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
KW	enzymatic assay; antibiotic target.
XX	
OS	Staphylococcus aureus.

XX	WO200294868-A2.
PN	
XX	
XX	
PD	
PD	28-NOV-2002.
XX	
XX	27-MAR-2002; 2002WO-IB002637.
PF	
PF	
XX	
XX	27-MAR-2001; 2001GB-00007661.
PR	

XX (CHIR-) CHIRON SPA.
XX
XX PI Massignani V, Mora M, Scarselli M;
XX

DR WPI; 2003-120786/11.
DR N-PSDB; ACF73443.

XX New *Staphylococcus aureus* protein, useful as a vaccine for treating or
PT preventing *Staphylococcal* infection, specifically an infection caused by
PT *S. aureus*, e.g. sepsis.

PS Claim 1; SEQ ID NO 2246; 49pp; English.

The invention relates to novel genes and encoded proteins from *Staphylococcus aureus*. A composition comprising the *S. aureus* protein, a nucleic acid encoding the protein, or an antibody to the protein, is useful as a pharmaceutical, particularly as a vaccine for treating or preventing infection due to *Staphylococcus bacteria*, specifically an infection caused by *S. aureus*. The composition is particularly useful for treating or preventing sepsis in a patient. The composition can also be used for diagnostics. The protein is also used in an assay for enzymatic studies and as a target for antibiotics. This sequence represents one of the novel *S. aureus* proteins of the invention.

xx Sequence 106 AA;
SQ

Query Match	24.4%;	Score	131.5;	DB	6;	Length	106;
Best Local Similarity	42.3%;	Pred. No.	1.3e-07;				
Matches	33;	Conservative	13;	Mismatches	31;	Indels	11;
						Gaps	1;

Qy	12	E P F K T L A H P A R I T E L L V E D R S V G E L L S D V G L E S S N L S Q O G L V L R A G V W A A R R D G N	71
		: : : : : : : :	
Db	18	E I F A L G D Y N R I R I M E L L S V S E A S V G H - I S H O L N L S O S N V S H O L K L J K S V H L V V A K R Q Q	76

QY 72 AMIYSIAAPDIAELLAVA 89

Db 77 SMIYSLDDIHVATMLKQA 94

RESULT 9
ABU28998

ABU28998
ID ABU28998 standard; protein; 95 AA.

AC ABU28998;

19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #14525.

Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Enterococcus faecalis.

PN WO200277183-A2.

PD 03-OCT-2002.

21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

FR 06-SEP-2001; 2001US-00348933.
PR 25-OCT-2001; 2001US-0342923P.

03-FEB-2002; 2002US-00072831
PR 06-MAR-2002; 2002US-0362699P.

PA (ELIT-) ELITRA PHARM INC.

PI	Wang L,	Zamudio C,	Malone C,	Haselbeck R,	Ohlsen KL,	Zyskind JW;
PI	Wall D,	Travick JD,	Carr GJ,	Yamamoto R,	Forsyth RA,	Xu HH;

DR WPI; 2003-029926/02.

DR N-PSDB; ACA32868.

PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX

PS Claim 25; SEQ ID NO 56922; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological

CC pathway in which a proliferation-required gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent

CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antisense nucleic acids are useful for

CC identifying proteins or screening for homologous nucleic acids required

CC for cellular proliferation to isolate candidate molecules for rational

CC drug discovery programs, or for screening homologous nucleic acids

CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of

CC the target prokaryotic essential genes. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 95 AA;

Query Match 23.7%; Score 127.5; DB 6; Length 95;

Best Local Similarity 36.7%; Pred No. 3.3e-07;

Matches 29; Conservative 20; Mismatches 29; Indels 1; Gaps 1;

QY 8 KKAEEFKTLAHPARIRILELVERDRSVGELLSSDVGLSSNLSQQLGVLRAGVVAAR 67

DB 8 KQVSQLYKVLSPDTRILULLKEGHNV-TAISQLGMEQSAVSHQLKLRDRSRVKAR 66

QY 68 RDCNAMIVSIAAPDAELL 86

DB 67 REGKTFYTLDDHHVDIL 85

RESULT 10

ABU40771

ID ABU40771 standard; protein; 107 AA.

XX

AC ABU40771;

XX

DT 19-JUN-2003 (first entry)

XX

DE Protein encoded by Prokaryotic essential gene #26298.

XX

KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX

OS Proteus sp.

XX

PN WO200277183-A2.

XX

PD 03-OCT-2002.

XX

PF 21-MAR-2002; 2002WO-US009107.

XX

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX

PA (ELIT-) ELITRA PHARM INC.

XX

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohleen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX

DR WPI: 2003-029926/02.

DR N-PSDB; ACA44641.

XX

PT New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX

PS Claim 25; SEQ ID NO 69695; 1766pp; English.

XX

CC The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological

CC pathway in which a proliferation-required gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent

CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antisense nucleic acids are useful for

CC identifying proteins or screening for homologous nucleic acids required

CC for cellular proliferation to isolate candidate molecules for rational

CC drug discovery programs, or for screening homologous nucleic acids

CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of

CC the target prokaryotic essential genes. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 107 AA;

Query Match 23.1%; Score 124.5; DB 6; Length 107;

Best Local Similarity 42.1%; Pred. No. 8.9e-07;

Matches 32; Conservative 12; Mismatches 31; Indels 1; Gaps 1;

QY 11 AEFKTLAHPARIRILELVERDRSVGELLSSDVGLSSNLSQQLGVLRAGVVAARDG 70

DB 21 SSMKLTGNGDRLLLCQLSGEKSVELES-LGIRPTLSQQLTVLRNEGLVNRDRD 79

QY 71 NAMYSTAAPDAELL 86

DB 80 KRIFYIADEKVIULL 95

RESULT 11

ADF06894

ID ADF06894 standard; protein; 121 AA.

XX

AC ADF06894;

XX

DT 12-FEB-2004 (first entry)

XX

DE Bacterial polypeptide #3007.

KW Proteus mirabilis infection; bacterial infection; antibacterial;
XX immunostimulant.
OS Proteus mirabilis.
XX
XX US6605709-B1.
XX
XX PD 12-AUG-2003.
XX
XX PF 05-APR-2000; 2000US-00543681.
XX
XX PR 09-APR-1999; 99US-0128706P.
XX
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX
XX PI Breton GL;
XX
XX DR WPI: 2003-895291/82.
XX
XX DR N-PSDB; ADF02722.
XX
XX PT New Proteus mirabilis polypeptides and polynucleotides, useful as
PT reagents for diagnosis of bacterial disease, as components of
PT antibacterial vaccines, as targets for antibacterial drugs, or as
PT biocontrol agents for plants.
XX
XX PS Disclosure; SEQ ID NO 7179; 870pp; English.
XX
XX CC The invention relates to new Proteus mirabilis polypeptides and
CC polynucleotides. The invention also relates to antibodies against the
CC polypeptides, methods for producing the polypeptides, a method of
CC generating vaccines for immunising an individual against P. mirabilis, a
CC method for evaluating a compound for the ability to bind a P. mirabilis
CC polypeptide and a method for screening test compounds for anti-bacterial
CC activity. The polypeptides and polynucleotides are useful as molecular
CC targets for diagnosing, preventing and treating pathological conditions
CC resulting from bacterial infection, as reagents for diagnosis of
CC bacterial diseases, as components of antibacterial vaccines, as targets
CC for antibacterial drugs or as bio-control agents for plants. This
CC sequence represents a Proteus mirabilis polypeptide of the invention.
XX
XX SQ Sequence 121 AA;
Query Match 23.1%; Score 124.5; DB 7; Length 121;
Best Local Similarity 42.1%; Pred. No. 1.1e-06;
Matches 32; Conservative 12; Mismatches 31; Indels 1; Gaps 1;
QY 11 AEFKTLAHPARIRILELLVERDRSVGELLSSDVGLESSLNLSQQLGVLRAGVVAARDG 70
Db 35 SSMKLTGNGDRLLLCQLSQGKSVSELES- LGIRQPTLSQQLVLRNGLVNRTRDG 93
QY 71 NAMIYSIAAPDIAELL 86
Db 94 KRIFYSIADKVLVLL 109
RESULT 12
ABU49708
ID ABU49708 standard; protein; 105 AA.
XX
AC ABU49708;
XX
XX 19-JUN-2003 (first entry)
XX
XX DE Protein encoded by Prokaryotic essential gene #35235.
XX
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Vibrio cholerae.
XX
XX PN WO200277183-A2.
XX
XX PD 03-OCT-2002.
XX

PF 21-MAR-2002; 2002WO-US009107.
XX
XX PR 21-MAR-2001; 2001US-00815242.
XX
XX PR 06-SEP-2001; 2001US-00948993.
XX
XX PR 25-OCT-2001; 2001US-0342923P.
XX
XX PR 08-FEB-2002; 2002US-00072851.
XX
XX PR 06-MAR-2002; 2002US-0362699P.
XX
XX PA (ELIT-) ELITRA PHARM INC.
XX
XX XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX DR WPI: 2003-029926/02.
XX
XX DR N-PSDB; ACA53578.
XX
XX PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX PS Claim 25; SEQ ID NO 77632; 1766pp; English.
XX
XX CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 105 AA;
Query Match 22.9%; Score 123.5; DB 6; Length 105;
Best Local Similarity 32.9%; Pred. No. 1.1e-06;
Matches 28; Conservative 16; Mismatches 40; Indels 1; Gaps 1;
QY 11 AEFKTLAHPARIRILELLVERDRSVGELLSSDVGLESSLNLSQQLGVLRAGVVAARDG 70
Db 22 AELLVMAHPERLMLVLCQLTQSEMVGVLQOGST-LSQSAFSQLTLVLRKHGIIQARKES 80
QY 71 NAMIYSIAAPDIAELLAVARKVLAR 95
Db 81 QQVFRYLRADSRITALIQLQNVCFR 105
RESULT 13
ABU18228
ID ABU18228 standard; protein; 96 AA.
XX

AC ABU18228;
 XX 19-JUN-2003 (first entry)
 XX Protein encoded by Prokaryotic essential gene #3755.
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KW Bacillus anthracis.
 OS
 XX
 PN WO2002771HJ-A2.
 XX 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haseelbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA22098.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 46152; 1766pp; English.
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 96 AA;
 SQ

Query Match 22.8%; Score 123; DB 6; Length 96;
 Best Local Similarity 33.7%; Pred. No. 1.2e-06;
 Matches 28; Conservative 18; Mismatches 37; Indels 0; Gaps 0;

OY 6 LYKLKAEFFKTLAHPARITLLELAVERDRSYCELLSSDVGLSSNLSQLGLVRAGVVA 65
 DB 11 LYESNAELLKVLAVPRLCIVKGLIERGPSNVSTMYTGLNMPQSTISQHLAKLSAKIVS 70
 OY 66 ARRDGNAMYSIAAPDIAELIAY 88
 DB 71 SERKGLIYYKVENETIIQLVRV 93
 RESULT 14
 ADB08342
 ID ADB08342 standard; protein; 107 AA.
 XX ADB08342;
 XX 20-NOV-2003 (first entry)
 XX Alloiooccus otitis antigenic protein SEQ ID NO:2282.
 DE Alloiooccus otitis; antigenic protein; immunogenic; immunisation;
 XX Alloiooccus otitis; Gram-positive bacterium; infection.
 KW gene therapy; Gram-positive bacterium; infection.
 XX Alloiooccus otitis.
 OS WO2003048304-A2.
 PN 12-JUN-2003.
 XX 25-NOV-2002; 2002WO-US036123.
 PF 29-NOV-2001; 2001US-0333777P.
 PR 18-NOV-2002; 2002US-036742P.
 XX (AMHP) WYETH HOLDINGS CORP.
 XX Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;
 PI WPI; 2003-505284/47.
 DR N-PSDB; ADB08341.
 XX New Alloiooccus otitis polynucleotides and polypeptides, useful for
 PT treating and diagnosing diseases, drug screening assays and monitoring of
 PT effects during drug clinical trials.
 XX Claim 33; SEQ ID NO 2282; 1019pp; English.
 XX The present invention describes an isolated polynucleotide (1) of
 CC Alloiooccus otitis genomic DNA, which encodes an antigenic protein.
 CC Alloiooccus otitis is a Gram-positive bacterium. Also described: (1)
 CC an isolated polypeptide that is encoded by the polynucleotide (1); (2) an
 CC expression vector comprising the novel isolated polynucleotide (1); its
 CC complement, degenerate variant or fragment; (3) a genetically engineered
 CC host cell, transfected, transformed or infected with the vector of (2);
 CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
 CC composition comprising the polypeptide, its complement, biological
 CC equivalent or fragment, or the polynucleotide that is comprised in the
 CC expression vector; (6) a pharmaceutical composition comprising the
 CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
 CC of the polypeptides of (1), their biological equivalent or fragment; (8)
 CC immunising against Alloiooccus otitis by administering to a host the
 CC immunogenic composition; (9) detecting and/or identifying Alloiooccus
 CC otitis in the biological sample; (10) a kit comprising a container
 CC containing the novel polynucleotide, its degenerate variant or fragment,
 CC or the antibody of (4); and (11) producing a polypeptide by culturing the
 CC genetically engineered host cell under conditions suitable to produce the
 CC polypeptide from the culture. (1) can be used in gene therapy. The
 CC polynucleotides, polypeptides, antibodies and compositions of the present
 CC invention can be used for treating and diagnosing diseases, drug
 CC screening assays and monitoring of effects during drug clinical trials.
 CC The polynucleotides are useful for expressing and detecting Alloiooccus
 CC otitis. The present sequence represents an Alloiooccus otitis
 CC antigen protein from the present invention.

```

CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 108 AA;
Query Match 22.2%; Score 119.5; DB 6; Length 108;
Best Local Similarity 33.3%; Pred. No. 3.6e-06;
Matches 26; Conservative 24; Mismatches 27; Indels 1; Gaps 1;
QY 15 KTLAHPARIRILELLVERDRSVGCELLSSDVGLESSNLSQLGLVLRAGVVAARRDGNAMI 74
Db 26 KAWANERRIQILCMLLDNLSVGE-LSSRLSELQSALSQHLAWLRDRDLVNTKREAOQTVF 84
QY 75 YSIAAPDIAELIAVARKV 92
Db 85 YTLSSTEVKAMIELLHRL 102
Search completed: October 15, 2005, 04:26:16
Job time : 31.3061 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 01:42:44 ; Search time 9.3599 Seconds
(without alignments)
909.197 Million cell updates/sec

Title: US-10-617-038-3
Perfect score: 539
Sequence: 1 VSEEPYKLBKAEFFKTLAHP.....RVLSDRVAVLEDLRAGGSAT 114

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124.5	23.1	121	4	US-09-543-681A-7179
2	119	22.1	113	4	US-09-710-279-1378
3	119	22.1	124	3	US-09-134-001C-4197
4	110	20.4	117	4	US-09-902-540-11185
5	96	17.8	348	4	US-09-252-991A-19898
6	93.5	17.3	103	4	US-09-328-352-7724
7	93	17.3	99	4	US-09-489-039A-9929
8	89	16.5	123	4	US-09-602-787A-560
9	89	16.5	123	4	US-09-603-208A-300
10	88.5	16.4	116	4	US-09-328-352-6454
11	85.5	15.9	142	4	US-09-252-991A-25372
12	83.5	15.5	110	3	US-09-134-001C-2984
13	83	15.4	110	4	US-09-328-352-8212
14	81	15.0	176	4	US-09-252-991A-24407
15	80.5	14.9	240	4	US-09-634-238-285
16	78.5	14.6	112	4	US-09-489-039A-11993
17	78	14.5	764	4	US-09-252-991A-31816
18	76	14.1	118	3	US-09-413-814-17
19	76	14.1	225	4	US-09-583-110-3071
20	76	14.1	231	4	US-09-107-433-3748
21	75.5	14.0	105	4	US-09-107-532A-6671
22	75.5	14.0	119	4	US-09-543-681A-7062
23	74.5	13.8	154	4	US-09-252-991A-17864
24	74.5	13.8	233	4	US-09-252-991A-27758
25	74.5	13.8	237	4	US-09-543-681A-4561
26	74	13.7	300	4	US-09-107-532A-4473
27	72.5	13.5	1233	4	US-09-252-991A-23237

28	72	13.4	120	4	US-09-107-532A-7083	Sequence 7083, Ap
29	72	13.4	312	4	US-09-489-039A-14078	Sequence 14078, A
30	72	13.4	410	4	US-09-902-540-12950	Sequence 12950, A
31	71	13.2	224	4	US-09-902-540-10058	Sequence 10058, A
32	71	13.2	770	4	US-09-489-039A-7872	Sequence 7872, Ap
33	70.5	13.1	996	4	US-09-252-991A-28596	Sequence 28596, A
34	70	13.0	712	4	US-09-252-991A-20911	Sequence 20911, A
35	69.5	12.9	112	4	US-09-252-991A-20911	Sequence 18866, A
36	69.5	12.9	659	4	US-09-902-540-16383	Sequence 16383, A
37	68.5	12.7	219	3	US-08-914-375C-3	Sequence 3, Appl1
38	68.5	12.7	416	4	US-09-252-991A-18835	Sequence 18835, A
39	68.5	12.7	600	4	US-09-252-991A-21703	Sequence 21703, A
40	68.5	12.7	663	4	US-09-248-796A-18398	Sequence 18398, A
41	68	12.6	152	4	US-09-902-540-13050	Sequence 13050, A
42	68	12.6	314	4	US-09-498-520A-34	Sequence 34, Appl
43	68	12.6	668	4	US-09-252-991A-21541	Sequence 21541, A
44	67.5	12.5	159	4	US-09-252-991A-18749	Sequence 18749, A
45	67.5	12.5	280	4	US-09-902-540-11498	Sequence 11498, A

ALIGNMENTS

RESULT 1

US-09-543-681A-7179
; Sequence 7179, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7179
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7179

Query Match 23.1%; Score 124.5; DB 4; Length 121;
Best Local Similarity 42.1%; Pred. No. 6.7e-08;
Matches 32; Conservative 12; Mismatches 31; Indels 1; Gaps 1;

QY	11	AEFFKTLAHPARIRLELLVERDSVGEILLSSDVGLSSNLSQLGLVLRAGVVAARRDG 70
DB	35	SSMLKTLGNGDRLLLCQLSQGKSVSELES-LGIQPTLSQQLTVLRLNEGLVNTTRDG 93
QY	71	NAMIYSIAAPDIAELL 86
DB	94	KRIFYSIADEKVLVLL 109

RESULT 2

US-09-710-279-1378
; Sequence 1378, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1378
; LENGTH: 113
; TYPE: PRT

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; US-09-710-279-1378

Query Match      22.1%; Score 119; DB 4; Length 113;
Best Local Similarity 37.4%; Pred. No. 3e-07;
Matches 34; Conservative 18; Mismatches 35; Indels 4; Gaps 3;

QY 1 VES--EPLYKLKAEFFKTLAHPARIRILELLVERDRSUGELLSSDVGLSSNLSQQLGVL 58
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 8 VESNQDTLIKV-THIFKALSDENVRIMEFLENGEASVGH-ISHSLNMTQSNVSHQLKLL 65
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

QY 59 VRAGVVAARDGNAMIYSTAAPDIAELLAVA 89
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 66 KSTHLVKSRQGSQSMIYSIDDIHVSTLLKQA 96
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

RESULT 3
US-09-134-001C-4197
; Sequence 4197, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4197
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4197

Query Match      22.1%; Score 119; DB 3; Length 124;
Best Local Similarity 37.4%; Pred. No. 3.4e-07;
Matches 34; Conservative 18; Mismatches 35; Indels 4; Gaps 3;

QY 1 VES--EPLYKLKAEFFKTLAHPARIRILELLVERDRSUGELLSSDVGLSSNLSQQLGVL 58
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 19 VESNQDTLIKV-THIFKALSDENVRIMEFLENGEASVGH-ISHSLNMTQSNVSHQLKLL 76
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

QY 59 VRAGVVAARDGNAMIYSTAAPDIAELLAVA 89
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 77 KSTHLVKSRQGSQSMIYSIDDIHVSTLLKQA 107
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

RESULT 4
US-09-902-540-11185
; Sequence 11185, Application US/0902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 11185
; LENGTH: 117
; TYPE: PRT
```

```
; ORGANISM: Myxococcus xanthus
US-09-902-540-11185

Query Match      20.4%; Score 110; DB 4; Length 117;
Best Local Similarity 29.6%; Pred. No. 4.3e-06;
Matches 32; Conservative 15; Mismatches 39; Indels 22; Gaps 2;

QY 14 PKTLAHPARIRILELLVERDRSUGEL---LSSDVGLSSNLSQQLGVLRRAGVVAARDG 70
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 4 FDLGPEPVRRIRILELLAEGRPSEGVETESIQAEFGISQPAVSQHLKVLURDSGFAEVRABG 63
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

QY 71 NAMIYSTAAPDIAELLA-----VARKVLARVLS 99
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 64 TRRLRYVDPAFLQQVDAWLEQFRVFWTPRLALATEVARGKRAKAD 111
   :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

RESULT 5
US-09-252-991A-19898
; Sequence 19898, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19898
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19898

Query Match      17.8%; Score 96; DB 4; Length 348;
Best Local Similarity 32.0%; Pred. No. 0.0011;
Matches 33; Conservative 14; Mismatches 54; Indels 2; Gaps 2;

QY 7 YKLKAEFFKTLAHPARIRILELLVERDRSUGELLSSDVGLSSNLSQQLGVLRRAGVVA 66
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 28 YDPLAALCKAAGDSLRLNVLRLANDSFGVLE-LAQIFAIGSQSGMSHHLKVLACAGLVAT 86
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

QY 67 RRDGNAMIYSTAAPDIAELLAVARKVLARVLSDRVAVLELRA 109
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 87 RREGNAIFRRSLP-LAERTGGAALHAALLEBVDERALPDDVRA 128
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

RESULT 6
US-09-328-352-7724
; Sequence 7724, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7724
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7724

Query Match      17.3%; Score 93.5; DB 4; Length 103;
Best Local Similarity 38.1%; Pred. No. 0.00043;
Matches 24; Conservative 13; Mismatches 25; Indels 1; Gaps 1;
```


RESULT 9

US-09-603-208A-300
; Sequence 300, Application US/09603208A
; Patent No. 6822084

; GENERAL INFORMATION:

; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; APPLICANT: Lee, Heung-Shick
; APPLICANT: Kim, Hyung-Joon

; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING STRESS,
; RESISTANCE AND TOLERANCE PROTEINS

; FILE REFERENCE: BGI-124CP

; CURRENT FILING DATE: 2000-06-23

; CURRENT APPLICATION NUMBER: US/09/603,208A

; PRIOR FILING DATE: 1999-06-25

; PRIOR APPLICATION NUMBER: 60/141031

; PRIOR FILING DATE: 1999-07-01

; PRIOR APPLICATION NUMBER: 60/142692

; PRIOR FILING DATE: 1999-08-27

; PRIOR APPLICATION NUMBER: DE 19930429.7

; PRIOR FILING DATE: 1999-07-01

; PRIOR APPLICATION NUMBER: DE 19931413.6

; PRIOR FILING DATE: 1999-07-08

; PRIOR APPLICATION NUMBER: DE 19931457.8

; PRIOR FILING DATE: 1999-07-08

; PRIOR APPLICATION NUMBER: DE 19931541.8

; PRIOR FILING DATE: 1999-07-08

; PRIOR APPLICATION NUMBER: DE 19932209.0

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932230.9

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932214.1

; PRIOR FILING DATE: 1999-07-14

; PRIOR APPLICATION NUMBER: DE 19940764.9

; PRIOR FILING DATE: 1999-08-27

; PRIOR APPLICATION NUMBER: DE 19941382.7

; PRIOR FILING DATE: 1999-08-31

; NUMBER OF SEQ ID NOS: 306

; SEQ ID NO 300

; LENGTH: 123

; TYPE: PRT

; ORGANISM: Corynebacterium glutamicum

US-09-603-208A-300

Query Match 16.5%; Score 89; DB 4; Length 123;

Best Local Similarity 31.4%; Pred. No. 0.002;

Matches 27; Conservative 17; Mismatches 34; Indels 8; Gaps 2;

QY 8 KLAEEFFKTLAHPARIRILELVERDRSVGEL----LSSDVGLESSLNSQQLGVLRAGV 63

DB 18 ELAVEVFGLLADATVRILAL-----RNSGELSVNHLADIVDKSPAIVSQHLRLRMARI 73

QY 64 VAARRDGNAMYISIAAPDIAELLAVA 89

DB 74 VSTRQEGQRFVFKLTNEHASQLVSDA 99

RESULT 10

US-09-328-352-6454

; Sequence 6454, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; RESISTANCE AND TOLERANCE PROTEINS

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 6454

; LENGTH: 116

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-6454

Query Match 16.4%; Score 88.5; DB 4; Length 116;

Best Local Similarity 30.1%; Pred. No. 0.0022;

Matches 22; Conservative 17; Mismatches 33; Indels 1; Gaps 1;

QY 10 KAEEFKTLAHPARIRILELVERDRSVGELSSDVGLESSLNSQQLGVLRAGVVAARRD 69

DB 11 QVDFKCLSDOTRNLKLVKNKONICVCELTEQLSQPKISRHLALLRTHGVLLDERK 70

QY 70 GNAMYISIAAPDI 82

DB 71 GQVYVYSL-NPDL 82

RESULT 11

US-09-252-991A-25372

; Sequence 25372, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 25372

; LENGTH: 142

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-25372

Query Match 15.9%; Score 85.5; DB 4; Length 142;

Best Local Similarity 29.8%; Pred. No. 0.0068;

Matches 34; Conservative 16; Mismatches 53; Indels 11; Gaps 4;

QY 1 VESEPLYKLKAEFFKTLAHPARIRILELVERDRSVGEL----LSSDVGLESSLNSQQLG 56

DB 21 IOEGPAMPSPAEPVFKCLADETRVATLLIVDQ---GELCVCELMCALADSQPKISRHLA 76

QY 57 VLRAGVVVAARRDGNAMYISI--AAPD-IAELLAVARKVLARVLSDRVAVLEDL 107

DB 77 QLRSGALLDRQGWVYVYRLNPTLPWAIHEVLTQLRANGDWLQADAARLDRM 130

RESULT 12

US-09-134-001C-2984

; Sequence 2984, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 2984

```
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2984

Query Match      15.5%; Score 83.5; DB 3; Length 110;
Best Local Similarity 30.7%; Pred. No. 0.0085;
Matches 23; Conservative 14; Mismatches 29; Indels 9; Gaps 2;

QY 7 YKLKAEFFKTLAHPARIRILELVERDRSVGELLSSDV-----GLESSNLSQQLGVLRAG 62
Db 8 YKLSAILKVLSPSLRLDILL-----SGELCACDILLAYFQFSPTLSHMKILVDNE 62

QY 63 VVAARDGNAMYSI 77
Db 63 LVSTRKDGTRMYRL 77

RESULT 13
US-09-328-352-8212
; Sequence 8212, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8212
; LENGTH: 110.
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-8212

Query Match      15.4%; Score 83; DB 4; Length 110;
Best Local Similarity 29.4%; Pred. No. 0.0099;
Matches 20; Conservative 15; Mismatches 33; Indels 0; Gaps 0;

QY 10 KAEFFKTLAHPARIRILELVERDRSVGELLSSDVGLSSNLSQQLGVLRAGVVAARD 69
Db 6 QVNFKCLSDETRLNIVTLIAENNELVCVDLTKQLSQPKISRHLALLRSSGLLDQRQ 65

QY 70 GNAMYSI 77
Db 66 SQWVYYSI 73

RESULT 14
US-09-252-991A-24407
; Sequence 24407, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24407
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24407

Query Match      15.0%; Score 81; DB 4; Length 176;
Best Local Similarity 32.4%; Pred. No. 0.034;
```

```
Matches 23; Conservative 8; Mismatches 28; Indels 12; Gaps 1;

QY 12 EFFEKTLAHPARIRILELVERDRSVGE-----LLSSDVGLESSNLSQQLGVLR 59
Db 83 ELIKALSHVPRDRMLRWLKEPEKYPVEQDHPFEIGVCAGKFDORTGLSQSTVSVHLATLQ 142

QY 60 RAGVVAARDG 70
Db 143 RAGLVTSSRVG 153

RESULT 15
US-09-634-238-285
; Sequence 285, Application US/09634238
; Patent No. 6544772
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Bloksberg, Leonard, N.
; APPLICANT: Lubbers, Mark W.
; APPLICANT: Dekker, James
; APPLICANT: Christensson, Anna C.
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul W.
; APPLICANT: Reid, Julian R.
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; FILE REFERENCE: 11000.1043U1
; CURRENT APPLICATION NUMBER: US/09/634,238
; CURRENT FILING DATE: 2000-08-08
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 285
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-09-634-238-285

Query Match      14.9%; Score 80.5; DB 4; Length 240;
Best Local Similarity 33.3%; Pred. No. 0.061;
Matches 29; Conservative 15; Mismatches 26; Indels 17; Gaps 5;

QY 37 GELLSSDVGLESSNLS-QQLGVLRAGVVAARDGNAMYISIAAPD-IAELLAVARKVLA 94
Db 99 GTVISSDLGLDKTKIEQLG---RAGKVTVTKDDTTIVDGAGSKDAIAERVNIIKKQIA 155

QY 95 RVLSD-----RVAVLEDLRAGGSA 113
Db 156 DTTSDFDREKLOERLAKL----AGGVA 178

Search completed: October 15, 2005, 03:24:06
Job time : 10.3599 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:14:10 : Search time 27.2619 Seconds
(without alignments)
1742.576 Million cell updates/sec

Title: US-10-617-038-3

Perfect score: 539

Sequence: 1 VESEPLYKLKAEFFKTLAHP.....RVLSDRVAVLEDLRAGGSAT 114

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	539	100.0	114	15	US-10-617-038-3
2	194	36.0	104	9	US-09-738-626-4472
3	141.5	26.3	218	15	US-10-282-122A-64593
4	135.5	25.1	226	15	US-10-282-122A-62549
5	135.5	25.1	226	15	US-10-282-122A-64367
6	127.5	23.7	95	15	US-10-282-122A-56922
7	124.5	23.1	107	15	US-10-282-122A-68695
8	123.5	22.9	105	15	US-10-282-122A-77632
9	123	22.8	96	15	US-10-282-122A-46152
10	122	22.6	107	15	US-10-501-282-2282
11	119.5	22.2	108	15	US-10-282-122A-77122

ALIGNMENTS

RESULT 1

US-10-617-038-3
; Sequence 3, Application US/10617038
; Publication No. US20040057963A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Peter
; APPLICANT: Rosenkrands, Ida
; APPLICANT: Stryhn, Anette
; TITLE OF INVENTION: Therapeutic TB Vaccine
; FILE REFERENCE: SS15AUSA
; CURRENT APPLICATION NUMBER: US/10/617,038
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: DK PA 2002 01098
; PRIOR FILING DATE: 2002-07-13
; PRIOR APPLICATION NUMBER: US 60/401,725
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-617-038-3

Query Match 100.0%; Score 539; DB 15; Length 114;
Best Local Similarity 100.0%; Pred. No. 1.7e-55;
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VESEPLYKLKAEFFKTLAHPARIRILELVERDRSVGELSSDVGLSSNLSQQLGVLRR 60
Db 1 VESEPLYKLKAEFFKTLAHPARIRILELVERDRSVGELSSDVGLSSNLSQQLGVLRR 60
QY 61 AGVVAARDGNAMIVSIAPDPAETAEILLAVARKVLARVLSDRVAVLEDLRAGGSAT 114


```
; SEQ ID NO 64367
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64367

Query Match      25.1%; Score 135.5; DB 15; Length 226;
Best Local Similarity 34.4%; Pred. No. 5e-07;
Matches 32; Conservative 20; Mismatches 40; Indels 1; Gaps 1;

QY   15 KTLAHPARIRILELVERDSVCGELSSDVGLESSLNSOQLGVLRRAGVVAAARDGNAMI 74
    :|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db   20 KALANGRRLOLDLLAQGERAV-EATATGMNLTITASANLQALKSGLVARREGTROY 78
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

QY   75 YSIAAPDIARELLAVARKVLARVLSDRVAVLEDL 107
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db   79 YRIAGEDVARLFALVGVVADEHLADVAVAADV 111
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

RESULT 6
US-10-282-122A-56922
; Sequence 56922, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56922
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-10-282-122A-56922

Query Match      23.7%; Score 127.5; DB 15; Length 95;
Best Local Similarity 36.7%; Pred. No. 1.4e-06;
Matches 29; Conservative 20; Mismatches 29; Indels 1; Gaps 1;
```

```

QY      8 KLKAEFFKTLAHPARIRILELLVERDRSVGELLSSDVGLSSNLSQQLGVLRAGVVAAR 67
Db      8 KQVSQLYKVLSDPTKRILLLLKEGHNV-TAISQQLGMEQSAVSHQKLRLDRSRVKAR 66

QY      68 RDGNAMIYSIAAPDIAELL 86
Db      67 REGKTFYTLDDHHVIDIL 85

RESULT 7
US-10-282-122A-68695
; Sequence 68695, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68695
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-10-282-122A-68695

Query Match      23.1%; Score 124.5; DB 15; Length 107;
Best Local Similarity 42.1%; Pred. No. 3.5e-06;
Matches 32; Conservative 12; Mismatches 31; Indels 1; Gaps 1;

QY      11 AEFFKTLAHPARIRILELLVERDRSVGELLSSDVGLSSNLSQQLGVLRAGVVAARDG 70
Db      21 SSMLKTLGNGDRLLLCQLSQGEKSVSELES-LGIQPTLSQQLTVLRNEGLVNRTRDG 79

QY      71 NAMIYSIAAPDIAELL 86
Db      80 KRIFYSIADKVLVLL 95

RESULT 8
US-10-282-122A-46152
; Sequence 46152, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith

```

```

US-10-282-122A-77632
; Sequence 77632, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77632
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-282-122A-77632

Query Match      22.9%; Score 123.5; DB 15; Length 105;
Best Local Similarity 32.9%; Pred. No. 4.4e-06;
Matches 28; Conservative 16; Mismatches 40; Indels 1; Gaps 1;

QY      11 AEFFKTLAHPARIRILELLVERDRSVGELLSSDVGLSSNLSQQLGVLRAGVVAARDG 70
Db      22 AELLRVMAHPERLAVLCQTQSENGVGLQOGST-LSQSAFQSHLTVLRKHGIIQARKS 80

QY      71 NAMIYSIAAPDIAELLAVARKVLAR 95
Db      81 QQVFYRLADSRITALIOSLQNVFCR 105

RESULT 9
US-10-282-122A-46152
; Sequence 46152, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith

```

```

; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2282
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Alloiococcus otitidis
US-10-501-282-2282

Query Match      22.6%; Score 122; DB 18; Length 107;
Best Local Similarity 36.0%; Pred. No. 6.7e-06;
Matches 32; Conservative 16; Mismatches 37; Indels 4; Gaps 2;

QY 11 ABEFKTLAHPAIRILELIVEDRSVGELSSDVGLESSLNLSQOLCVLRAGVVAARRDG 70
DB 23 AQTFKALSDPTRIRILYLAAHECSVNE-IADTLDLKQSTVSHQLRNKLSNLVKKRREA 81
QY 71 NAMIVSIAAPDIAELLAVARKVLRVLSLD 99
DB 82 TSLIYS---PDDQHVLDSLKAEIHALHD 107

RESULT 11
US-10-282-122A-77122
; Sequence 77122, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liansu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 77122
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-10-282-122A-77122

Query Match      22.2%; Score 119.5; DB 15; Length 108;
Best Local Similarity 33.3%; Pred. No. 1.3e-05;
Matches 26; Conservative 24; Mismatches 27; Indels 1; Gaps 1;

```

```
Qy      15 KTLAHPARIRILELLVERDRSVGELLSSDVGLESSNLSQQGLVLRAGVAARDGNAMI   74  
       |:|::|||:::|||::|||::|||::|||::|||::|||::|||::|||::|||:  
Db      26 KAMANERRQLTCLMDNELSNGE-LSSRLELSQLSHLAWLRRDDGLVNTRKEAQTVF    84  
  
Qy      75 YSIAAPDAIELAVARKV          92  
       ||:::::::::::  
Db      85 YTSLSTEVKAMIEILLHRL        102  
  
RESULT 12  
US-10-724-972A-3812  
; Sequence 3812, Application US/10724972A  
; Publication No. US20040147734A1  
; GENERAL INFORMATION:  
; APPLICANT: Doucette-Stamm, Lynn  
; APPLICANT: Bush, David  
; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: PATHOJ3-16  
; CURRENT APPLICATION NUMBER: US/10/724,972A  
; CURRENT FILING DATE: 2003-12-01  
; PRIOR APPLICATION NUMBER: 09/450,969  
; PRIOR FILING DATE: 1999-11-29  
; PRIOR APPLICATION NUMBER: 09/134,001  
; PRIOR FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 7544  
; SEQ ID NO 3812  
; LENGTH: 124  
; TYPE: PRT  
; ORGANISM: S.epidermidis  
US-10-724-972A-3812
```

```
Query Match           22.1%; Score 119; DB 18; Length 124;  
Best Local Similarity 37.4%; Pred. No. 1.8e-05;  
Matches 34; Conservative 18; Mismatches 35; Indels 4; Gaps 3;
```

```
Qy      1 VES--EPYLKKAEEFFKTLAHPARIIRILELVDRSRVSGELSSDVGLESSLNSOOLGVL   58  
       ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db      19 VESNQDTLIK-VTHFKALSDFNVRINEFEFLENGASVGH-ISHLNTQTQNVSHQLKL    76  
  
Qy      59 RRGVVAAARRDGNMIYSIAAPDAIELAVA          89  
       :::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Db      77 KSTHLVKSKRCQSQMYSIIDDIVHTLKQA         107  
  
RESULT 13  
US-10-282-122A-49733  
; Sequence 49733, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21
```


THIS PAGE BLANK (00870)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:59:30 ; Search time 6.3611 Seconds
(without alignments)
1724.342 Million cell updates/sec

Title: US-10-617-038-3
Perfect score: 539
Sequence: 1 VSEEPYKLEKAEFFKTLAHP.....RVLSDRVAVLEDLRAGGSAT 114

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues 283416
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	536	99.4	114	2 B70850	probable transcript
2	315.5	58.5	117	2 T36657	probable transcript
3	151.5	28.1	115	2 C95333	probable transcript
4	145.5	27.0	117	2 A13610	transcription regu
5	144.5	26.8	114	2 D82766	transcription regu
6	142.5	26.4	116	2 A12982	transcription regu
7	142.5	26.4	117	2 F98300	transcription regu
8	141.5	26.3	218	2 E70818	probable mercury r
9	140.5	26.1	125	2 S30170	mercury resistance
10	135.5	25.1	226	2 E70526	hypothetical prote
11	135	25.0	115	2 E90443	hypothetical prote
12	134.5	25.0	123	2 H87603	transcription regu
13	131.5	24.4	106	2 T44366	hypothetical prote
14	128.5	23.8	122	2 S19675	nolR protein - Rhi
15	124.5	23.1	108	2 H72306	heavy metal resist
16	123.5	22.9	105	2 H82433	transcription regu
17	120.5	22.4	106	2 AE2856	transcription regu
18	120.5	22.4	119	2 B97633	nolR protein (A00
19	119.5	22.2	108	2 S37313	transcription acti
20	119	22.1	125	2 A82293	transcription regu
21	119	22.1	151	2 H87624	transcription regu
22	118.5	22.0	130	2 C70811	hypothetical prote
23	116.5	21.6	107	2 E69930	transcription regu
24	115.5	21.4	109	2 G75373	transcription regu
25	115	21.3	101	2 AG1761	transcription regu
26	114.5	21.2	114	2 AH2437	transcription regu
27	114.5	21.2	118	2 D75141	transcription repr
28	114.5	21.2	120	2 AH2273	transcription regu
29	114.5	21.2	122	2 S31197	smtB protein - Syn

30	114	21.2	101	2 AE1386	transcription regu
31	113.5	21.1	118	2 H70757	probable transcript
32	112.5	20.9	110	2 F95983	probable transcript
33	112.5	20.9	115	2 AH3305	modulation protei
34	112.5	20.9	140	2 C87012	probable ArsR-fam
35	111.5	20.7	108	2 AG2151	transcription regu
36	111	20.6	120	2 A70799	probable transcript
37	109	20.2	89	2 D64465	hypothetical prote
38	109	20.2	120	2 T35946	probable regulator
39	109	20.2	196	2 A11936	transcription regu
40	108.5	20.1	122	2 B69220	transcription regu
41	108.5	20.1	368	2 T36075	probable transcript
42	106.5	19.8	434	2 G70933	probable transcript
43	105.5	19.6	132	2 S77014	hypothetical repr
44	104.5	19.4	185	2 B71208	hypothetical prote
45	103.5	19.2	89	2 C72286	transcription regu

ALIGNMENTS

RESULT 1

B70850
probable transcription regulator Rv0081 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: B70850
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: B70850
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-114 <COL>
A;Cross-references: UNIPROT:O53626; GB:AL021428; GB:AL123456; NID:g3261514; PIDN:CAA1626; C;Genetics:
A;Experimental source: strain H37RV
A;Gene: Rv0081

Query Match 99.4%; Score 536; DB 2; Length 114;
Best Local Similarity 99.1%; Pred. No. 2.9e-43;
Matches 113; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSEEPYKLEKAEFFKTLAHPARIRILELVERDRSVGELLSSDVGLSSNLSQLGVLR 60

Db 1 MESEPYKLEKAEFFKTLAHPARIRILELVERDRSVGELLSSDVGLSSNLSQLGVLR 60

Qy 61 AGVVAARRDGNMIYIAAPDIAELLAVARKVLARVLSDRVAVLEDLRAGGSAT 114

Db 61 AGVVAARRDGNMIYIAAPDIAELLAVARKVLARVLSDRVAVLEDLRAGGSAT 114

RESULT 2

T36657
probable transcription regulator - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T36657
R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, June 1999
A;Reference number: Z21610
A;Accession: T36657
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-117 <OLI>
A;Cross-references: UNIPROT:Q9X904; EMBL:AL078610; PIDN:CAB44425.1; GSPDB:GN00070; SCOD A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCODB:SCH35.02

```

Query Match      58.5%; Score 315.5; DB 2; Length 117;
Best Local Similarity 58.2%; Pred. No. 1.3e-22;
Matches 64; Conservative 23; Mismatches 22; Indels 1; Gaps 1;

QY 3 SBPLYKLAFFKTLAHPARIRILELLVERDRSVGELSSDVGLESSLNSQQLGVLRRAG 62
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2 STPLQKLAFFKTLAHPARIRIVLELLSERHVAEML-PEVGIEPAHLSQQLAVLRAN 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 63 VVAARRDGNAMTYSTAAPDIAELLAVARKVLARVLSDRVAVLEDRAGS 112
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 LVVSRKRGSTVYSLTSPQVALLRVARGILSGVLQAGAEALLADLRAGS 110
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 3
C95333
Probable transcription regulator [imported] - Sinorhizobium meliloti (strain 1021) megap
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: C95333
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
.; Kaiman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: C95333
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-115 <KUR>
A:Cross-references: UNIPROT:Q922B9; GB:AE006469; PIDN:AAK65229.1; PID:g14523678; GSPDB:G
A:Experimental source: strain 1021, megaplasmid pSymA
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hymen, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kaiman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
A:Genetics:
A:Gene: Sma1056
A:Genome: plasmid

Query Match      28.1%; Score 151.5; DB 2; Length 115;
Best Local Similarity 40.6%; Pred. No. 2.7e-07;
Matches 39; Conservative 13; Mismatches 35; Indels 9; Gaps 2;

QY 11 ABEFFKTLAHPARIRILELLVERDRSVGELSSDVGLESSLNSQQLGVLRRAGVAAARDG 70
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 28 AKLLKTLAHPARIMVCTLVQGEYSVGE-LEEKVDVQHPHLSQHLTVLRGSGIVQTRRDG 86
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 71 NAMIYSIAAPDIAELLAVARKVLARVLSDRVAVLEDR 106
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 87 KQIFVRLTGEKAARLIA-----ALYDIFCVKED 114
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 4
A13610
Transcription regulator, arsr family BMEII0810 [imported] - Brucella melitensis (strain
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: A13610
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AD3252; PMID:11756688
A:Accession: A13610
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-117 <KUR>
A:Cross-references: UNIPROT:Q8YBS6; GB:AE008918; PIDN:AAU54052.1; PID:g17985006; GSPDB:G

```

```

A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEII0810
A:Map position: 11

Query Match      27.0%; Score 145.5; DB 2; Length 117;
Best Local Similarity 43.4%; Pred. No. 1e-06;
Matches 33; Conservative 12; Mismatches 30; Indels 1; Gaps 1;

QY 11 ABEFFKTLAHPARIRILELLVERDRSVGELSSDVGLESSLNSQQLGVLRRAGVAAARDG 70
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 29 AELLKTLAHPARIMVCTLVQGEYSVGE-LEEMLGITQPTLSQQLGVLREADIVETRREA 87
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 71 NAMIYSIAAPDIAELL 86
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 88 KOIYRLVETVETKAALKI 103
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
D82766
Transcription regulator Arsr family XF0767 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: D82766
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequenc
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: For a complete list of authors see reference number A59328 below
A:Accession: D82766
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-114 <SIM>
A:Cross-references: UNIPROT:Q9PFB1; GB:AE003917; GB:AE003949; NID:g9105648; PIDN:AAF83573
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarunga, R.; AJ
B:Riones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramao, E.E.; Laigre
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.;
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Sancelli, R.V.; Sawaraki
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
A:Genetics:
A:Gene: XF0767

Query Match      26.8%; Score 144.5; DB 2; Length 114;
Best Local Similarity 42.1%; Pred. No. 1.2e-06;
Matches 32; Conservative 14; Mismatches 29; Indels 1; Gaps 1;

QY 11 ABEFFKTLAHPARIRILELLVERDRSVGELSSDVGLESSLNSQQLGVLRRAGVAAARDG 70
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26 ANLLKTLAHPARIMVCTLVQGEYSVGE-LEQQIGIGQFTLSQQLGVLRESGIVETRRNI 84
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 71 NAMIYSIAAPDIAELL 86
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 85 KQIFVRLTETKAQALV 100
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 6
A12982
Transcription regulator, Arsr family Atu3466 [imported] - Agrobacterium tumefaciens (stra
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: A12982
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gilliet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell

```


A::title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A::Reference number: A70500; MUID:98295987; PMID:9634230
A::Accession: E70818
A::Status: preliminary; nucleic acid sequence not shown; translation not shown
A::Molecule type: DNA
A::Residues: 1-218 <COL>
A::Cross-references: UNIPROT:O53921; GB:AL022002; GB:AL123456; NID:g3261544; PIDN:CAA17604
A::Experimental source: strain H37Rv
C::Genetics:
A::Gene: RV1674c

Query Match 26.3%; Score 141.5; DB 2; Length 218;
Best Local Similarity 39.3%; Pred. No. 4.8e-06;
Matches 33; Conservative 21; Mismatches 28; Indels 1; Gaps 1;

QY 15 KTLAHPARIRILELLVDRDSVGELSSDVGLESSLQQGLGVLRAGVVAARRDCGNAMI 74
 : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 17 QALLSSGHRLELDDLVOGRSV-DALARASGLTANASQHLLQLRRAGLVTSRRDGKRV 75
 : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY 75 YSTAAPDIABELLAVARKVLARVLS 98
 | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
DB 76 YALSDPQVDWVRVRAVRAERNLA 99
 | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :

RESULT 9
S30170
mercury resistance operon regulator merR - Streptomyces lividans
N::Alternate names: mercuric regulator merR
C::Species: Streptomyces lividans
C::Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text_change 09-Jul-2004
C::Accession: S30170; S70625; S23610
R::Sedlmeier, R.; Altenbuchner, J.
Mol. Gen. Genet. 236, 76-85, 1992
A::Title: Cloning and DNA sequence analysis of the mercury resistance genes of Streptomyces
A::Reference number: S30168; MUID:93156687; PMID:1494353
A::Accession: S30170
A::Molecule type: DNA
A::Residues: 1-125 <ALT>
A::Cross-references: UNIPROT:P30346; EMBL:X65467; NID:g47180; PIDN:CAA46462.1; PID:g47183
A::Experimental source: strain 1326
R::Brunker, P.; Rother, D.; Sedlmeier, R.; Klein, J.; Mattes, R.; Altenbuchner, J.
Mol. Gen. Genet. 251, 307-315, 1996
A::Title: Regulation of the operon responsible for broad-spectrum mercury resistance in St
A::Reference number: S70625; MUID:96262183; PMID:8676873
A::Accession: S70625
A::Status: not compared with conceptual translation
A::Molecule type: DNA
A::Residues: 1-40 <BRU>
A::Experimental source: strain 1326
C::Genetics:
A::Gene: merR
C::Function:
A::Description: represses transcription of mercury resistance operon in the absence of mer
A::Note: DNA binding of MerR is prevented by mercuric ions and phenylmercuric acetate; mer
C::Superfamily: arsenical resistance operon repressor
C::Keywords: DNA binding; transcription regulation

Query Match 26.1%; Score 140.5; DB 2; Length 125;
Best Local Similarity 33.7%; Pred. No. 3.2e-06;
Matches 32; Conservative 21; Mismatches 37; Indels 5; Gaps 2;

QY 11 AEFFKLTAHPARIRILELLVDRDSVGELSSDVGLESSLQQGLGVLRAGVVAARRDG 70
 | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
DB 24 ARFFRALADTRLKLLQFILRGERTSAECV-EHAGISQPRVSVHLSCSLVDCGCVVSARRDG 82
 | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :

QY 71 NAMIYSTAAPDIAELLAVARKVLARVLSDRVAVLE 105
 | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :
DB 83 KKLRYSGDPRVADLVMLARCLAA---DNAALD 113
 | : : | : : | : : | : : | : : | : : | : : | : : | : : | : : | : :

RESULT 10
E70526
hypothetical protein RV0324 - Mycobacterium tuberculosis (strain H37Rv)

C:Species: *Mycobacterium tuberculosis*
C:date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: E70526
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Cole, S.T.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sultston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: E70526
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-226 <COL>
A:Cross-references: UNIPROT:O08446; GB:Z96800; GB:AL123456; NID:g3261800; PIDN:CAB09589
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: RV0324

Query Match	25.1%	Score 135.5	DB 2	Length 226
Best Local Similarity	34.4%	Pred. No. 1.8e-05		
Matches	32	Conservative 20	Mismatches 40	Indels 1
Gaps	1			

Qy	15	KTIAHPARIILELLVDRDSVGGELSSDVGLESSNLSQQGLGVLRRAQVVAARDGNAMI	74
Db	20	KALANGRRQLDILLAQERAV-EAIAATATGMNLTASANLQALKSGELVBARREGTQY	78

Qy	75	YSIAAPDIAELLAVARKVLARVLSDRVAVLEDL	107
Db	79	YRIAGEDVARLFALVQVVADEHLDAVAAADV	111

RESULT 11
E90443
hypothetical protein mefR [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: E90443
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awavez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.

Query Match	25.0%	Score 135;	DB 2;	Length 115;
Best Local Similarity	31.3%	Pred. No. 9.5e-06;		
Matches	31;	Conservative 21;	Mismatches 43;	Indels 4;
Gaps	1;			
Qy	4	EPLYKKAEPFKYLAHPARIRLELLVERDRSVEGLLESSNLSQOLGVLRAGV	63	
Db	2	EPLTNELESFSAADGTRLRIVFLDKGQATVDEISKSLGKSQLISHHMACLRNCGI	61	
Qy	64	VAARDGNAMITYSTAAPDIAELLAVA----	RKVLARVLS	98
Db	62	VKVRKDGKFSYISTPTEIIEIKLSINHVKKYSQSILS	100	

RESULT 12
H87603
transcription regulator, *AraG* family [imported] - *Caulobacter crescentus*
C:Species: *Caulobacter crescentus*
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: H87603
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoev, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
...

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
 A:Reference number: A87249; MUID:21173698; PMID:11259647
 A:Accession: H87603
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-123 <STO>
 A:Cross-references: UNIPROT:Q9A4H0; GB:AE005673; NID:gl3424478; PIDN:AAK24828.1; GSPDB:GN
 C:Genetics:
 A:Gene: CC2864

Query Match	25.0%	Score 134.5;	DB 2;	Length 123;
Best Local Similarity	41.5%;	Pred. No. 1.le-05;		
Matches	34;	Conservative 15;	Mismatches 32;	Indels 1; Gaps 1;
Qy	11	AEEFKTLHAPARIRIELLWVRDRSRVGGELSSDVGLESSLNLSQQLGVLRAGVVVAARDG	70	
	:	:	:	:
	:	:	:	:
Db	23	AKLLRALSNERRIMILCQAEGERSVGQ-IQPLVLGSQSALSQHIAVLRLDEGVVAARKEA	81	
	:	:	:	:
Qy	71	NAMIYSIAPDIALLAVARKV	92	
	:	:	:	:
Db	82	TTVMYRIADPAALKVATLAEI	103	

RESULT 13
T44366
Hypothetical protein CzrA [imported] - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T44366
R:Kuroda, M.; Hayashi, H.; Ohta, T.
Microbiol. Immunol. 43, 115-125, 1999
A:title: Chromosome-determined zinc responsible operon czr in Staphylococcus
A:reference number: Z22754; MUID:9244271; PMID:1029265
A:Accession: T44366
A:Status: preliminary; translated from GB/EMBL/DBDJB
A:Molecule type: DNA
A:Residues: 1-106 <RUR>
A:Cross-references: UNIPROT:O85142; EMBL:AB016431; PIDN:BAA36687.1
A:Experimental source: strain 912
C:Genetics: C
A:Note: CzrA

Query Match 24.4%; Score 131.5; DB 2; Length 106;
Best Local Similarity 42.3%; Pred. NO. 1.8e-05;
Matches 33; Conservative 13; Mismatches 31; Indels 1; Gaps 1;

```

Qy      72 AMIYSIAAPDIAELLAVA 89
       :|||:      :|:|
Db      77 SMYSLDDIHVATMLKQA 94

RESULT 14
S19675
S19675      nolr protein - Rhizobium meliloti
C/Species: Rhizobium meliloti
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C/Accession: S19675
R:Kondorosi, E.; Pierre, M.; Cron, M.; Haumann, U.; Buire, M.; Hoffmann, B.; Schell, J.;
J. Mol. Biol. 222, 885-896, 1991
A/Title: Identification of NolR, a negative transacting factor controlling the nod regulon
A/Reference number: S19675; PMID:92106347; PMID:1840615
A/Accession: S19675
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-122 <KON>
A/Cross-references: UNIPROT:P28267; EMBL:X59050; NID:g46314; PIDN:CAA41777.1; PID:g46315
C/Superfamily: arsenical resistance operon repressor

```

C;Keywords: DNA binding; transcription regulation

Query Match 23.8%; Score 128.5; DB 2; Length 122;
Best Local Similarity 34.0%; Pred. No. 4.1e-05;
Matches 34; Conservative 18; Mismatches 35; Indels 13; Gaps 2;

OY 8 KLKAEFFKTLAHPARIRILELLVERDRSVGELLSSDVGLSSNLSQOLGVLRRAGVVAAR 67
DB 21 ETAAAGFLSAMANKRLIILDSLVKEEMAVG-ALAHKVGLSQSALSQHLSKLRAQNLVSTR 79

OY 68 RQGNAMIYSIAAPDIAELL-----AVARKVLAR 95

DB 80 RDAQTIYSSSDAVLKILGALSIDIYDDTDAVEEKPLVR 119

RESULT 15

H72306
heavy metal resistance transcription regulator - Thermotoga maritima (strain MSB8)

C;Species: Thermotoga maritima

C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004

C;Accession: H72306

R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.

Nature 399, 323-329, 1999

A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571

A;Accession: H72306

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-108 <ARN>

A;Cross-references: UNIPROT:Q9X093; GB:AE001762; GB:AE000512; NID:g4981535; PIDN:AAD3608

A;Experimental source: strain MSB8

C;Genetics:

A;Gene: TM0998

C;Superfamily: arsenical resistance operon repressor

Query Match 23.1%; Score 124.5; DB 2; Length 108;
Best Local Similarity 36.0%; Pred. No. 8.5e-05;
Matches 31; Conservative 20; Mismatches 28; Indels 7; Gaps 2;

OY 12 EEFKTLAHPARIRILELLVERDRSVGELLSSDVGLSSNLSQOLGVLRRAGVVAARRDGN 71

DB 5 ELFHILSNETRLKILTLLEKEMCVCOILAS-IGTQPNISQHLHLVKNHGIVKSRREDS 63

OY 72 AMIYSI-----AAPDIAELLAVARK 91

DB 64 FVYISIDERFLEKYPPFLITILERARK 89

Search completed: October 15, 2005, 04:38:20
Job time : 7.3611 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:24:25 ; Search time 25.5353 Seconds
(without alignments)
2286.133 Million cell updates/sec

Title: US-10-617-038-3

Perfect score: 539

Sequence: 1 VSEEPYKLUKABFFKTLAHP.....RVLSDRVAVLDRAGGSAT 114

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Uniprot_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	536	99.4	114	2	O53626	053626 mycobacteri
2	536	99.4	114	2	Q7U2W1	Q7u2w1 mycobacteri
3	315.5	58.5	117	2	Q9X304	Q9x304 streptomyc
4	251.5	46.7	113	2	Q9PB31	Q9pb31 streptomyc
5	220	40.8	133	2	Q8PR22	Q8pr22 corynebacte
6	194	36.0	104	2	Q8NS13	Q8ns13 corynebacte
7	151.5	28.1	115	2	Q92ZB9	Q92zb9 rhizobium m
8	150.5	27.9	99	2	Q67SE8	Q67se8 symbiobacte
9	146.5	27.2	112	2	Q98J11	Q98j11 rhizobium l
10	145.5	27.0	107	2	Q8FWJ7	Q8fwj7 bruceella su
11	145.5	27.0	117	2	Q8VBS6	Q8vbs6 bruceella me
12	145.5	27.0	149	2	Q9LCV3	Q9lcv3 paracoccus
13	144.5	26.8	91	2	Q9K148	Q9k148 aphanocethe
14	144.5	26.8	114	2	Q87AD5	Q87ad5 xylella fas
15	144.5	26.8	114	2	Q9FPB1	Q9fpb1 xylella fas
16	144.5	26.8	124	2	Q6N1D2	Q6n1d2 rhodopseudo
17	143.5	26.6	113	2	Q6NB11	Q6nb11 rhodopseudo
18	142.5	26.4	116	2	Q8UAA8	Q8uaa8 agrobacteri
19	142.5	26.4	117	2	Q7CSJ3	Q7csj3 agrobacteri
20	141.5	26.3	112	2	Q6ZEU5	Q6zeu5 synecocyst
21	141.5	26.3	125	2	Q9L7K9	Q9l7k9 streptomyc
22	141.5	26.3	218	2	O53921	O53921 mycobacteri
23	141.5	26.3	218	2	Q7TZT6	Q7tzt6 mycobacteri
24	140.5	26.1	125	1	MERR_STRLI	P30346 streptomyc
25	140	26.0	108	2	Q97ZQ8	Q97zq8 sulfobobus
26	139.5	25.9	106	2	Q7P1X8	Q7p1x8 chromobacte
27	139.5	25.9	115	2	Q7UNAB	Q7uena8 rhodopirell
28	137.5	25.5	120	2	Q7UBN9	Q7uen9 rhodopirell
29	137.5	25.5	121	2	O8XZ61	O8xz61 ralstonia s
30	137	25.4	106	2	O89166	O89166 bradyrhizob
31	136	25.2	226	2	Q9Z1F6	Q9z1f6 streptomyc

32 135.5 25.1 226 2 Q7DA08
33 135.5 25.1 226 2 O08446
34 135.5 25.1 226 2 Q7U294
35 135 25.0 115 2 Q97VE0
36 134.5 25.0 116 2 O89PH4
37 134.5 25.0 123 2 Q9A4H0
38 134 24.9 129 2 Q67MF9
39 133.5 24.8 105 2 Q92FP93
40 132.5 24.6 118 2 Q8JTD2
41 132.5 24.6 118 2 O84FJ1
42 131.5 24.4 97 2 O7NSR9
43 131.5 24.4 106 2 O85142
44 131.5 24.4 106 2 Q7A0B0
45 131.5 24.4 106 2 Q7A2M9

ALIGNMENTS

RESULT 1
O53626 PRELIMINARY; PRT; 114 AA.
ID AC O53626 Q7DAH4;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN (Transcriptional
DE regulator, Arsr family).
GN OrderedLocusNames=MT0088, Rv0081;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J.J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrall B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
CC -!- SIMILARITY: Contains 1 HTH arsr-type DNA-binding domain.
DR EMBL; BX842572; CAA16262.1; -;
DR EMBL; AF000516; AAK44313.1; -;
DR PIR; B70850; B70850.
DR TIGR; MT0088; -;
DR TubercuList; Rv0081; -;
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001845; HTH_ArsR.
DR InterPro; IPR009058; Wing_hix_DNA_bnd.

Q7da08 mycobacteri
O08446 mycobacteri
Q7u294 mycobacteri
Q97ve0 sulfobobus
O89ph4 bradyrhizob
Q9a4h0 caulobacter
Q67mf9 symbiobacte
Q92fp93 rhizobium m
Q8jtd2 rhizobium f
Q84fj1 rhizobium s
O7nsr9 chromobacte
O85142 staphylococ
Q7a0b0 staphylococ
Q7a2m9 staphylococ

```

DR Pfam: PF01022; HTH_5; 1.
DR PRINTS: PR00778; HTHARSR.
DR SMART: SM00418; HTH_ARSR; 1.
KW DNA-binding; Transcription; Transcription regulation.
SQ SEQUENCE 114 AA; 12356 MW; AF3C88714DCC1B4F CRC64;

Query Match
Best Local Similarity 99.1%; Score 536; DB 2; Length 114;
Matches 113; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VESEPLYLKAEFFKTLAHPARIRILELLVERDRSVGELSSDVGLESSNLSQQGLVLR 60
Db 1 MESEPLYLKAEFFKTLAHPARIRILELLVERDRSVGELSSDVGLESSNLSQQGLVLR 60

QY 61 AGVVAARDGNAMYSIAAPDIAELLAVARKVLRLSDRVAVLEDLRAGGSAT 114
Db 61 AGVVAARDGNAMYSIAAPDIAELLAVARKVLRLSDRVAVLEDLRAGGSAT 114

RESULT 2
QYU2W1 PRELIMINARY; PRT; 114 AA.
AC QYU2W1;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DE PROBABLE TRANSCRIPTIONAL REGULATORY PROTEIN.
GN OrderedLocustNames=Mb0084;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempe C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL: BX248334; CAD92946.1; -.
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR001845; HTH_ArsR.
DR Pfam: PF01022; HTH_5; 1.
KW Complete proteome.
SQ SEQUENCE 114 AA; 12356 MW; AF3C88714DCC1B4F CRC64;

Query Match
Best Local Similarity 99.1%; Score 536; DB 2; Length 114;
Matches 113; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VESEPLYLKAEFFKTLAHPARIRILELLVERDRSVGELSSDVGLESSNLSQQGLVLR 60
Db 1 MESEPLYLKAEFFKTLAHPARIRILELLVERDRSVGELSSDVGLESSNLSQQGLVLR 60

QY 61 AGVVAARDGNAMYSIAAPDIAELLAVARKVLRLSDRVAVLEDLRAGGSAT 114
Db 61 AGVVAARDGNAMYSIAAPDIAELLAVARKVLRLSDRVAVLEDLRAGGSAT 114

RESULT 3
QYX904 PRELIMINARY; PRT; 117 AA.
AC QYX904;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DE Putative transcriptional regulator.

```

```

GN ORFNames=SCH35.02;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinovich E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RL Nature 417:141-147(2002).
CC -!- SIMILARITY: Contains 1 HTH arsr-type DNA-binding domain.
DR EMBL: AL939117; CAB44425.1; -.
DR PIR: T36657; T36657.
DR HSSP: P30340; 1SWT.
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR001845; HTH_ArsR.
DR InterPro: IPR009058; Wing_hlx_DNA_bnd.
DR Pfam: PF01022; HTH_5; 1.
DR PRINTS: PR00778; HTHARSR.
DR SMART: SM00418; HTH_ARSR; 1.
KW Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 117 AA; 12674 MW; D5FB3CA26B4E279E CRC64;

Query Match
Best Local Similarity 58.5%; Score 315.5; DB 2; Length 117;
Matches 64; Conservative 23; Mismatches 22; Indels 1; Gaps 1;

QY 3 SEPLYLKAEFFKTLAHPARIRILELLVERDRSVGELSSDVGLESSNLSQQGLVLR 62
Db 2 STPLYLKAEFFKTLGHPARIRVLELLSEREHAEML-PEVGIEPAHLSQQAVLR 60

QY 63 VVAARRDGNAMYSIAAPDIAELLAVARKVLRLSDRVAVLEDLRAGGS 112
Db 61 LWSRKEGSTVYVSLTSPQVAELLRVARGILSGVLGQAEALLADLRAGQS 110

RESULT 4
QYFBJ1 PRELIMINARY; PRT; 113 AA.
AC QYFBJ1;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Metal-dependent regulatory protein.
OS Streptomyces verticillus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=29309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC15003;
RX MEDLINE=20389599; PubMed=10930733;
RA Du L., Chen M., Sanchez C., Shen B.;
RT "An oxidation domain in the BimIII non-ribosomal peptide synthetase
RT probably catalyzing thiazole formation in the biosynthesis of the
RT anti-tumor drug bleomycin in Streptomyces verticillus ATCC15003."
RL FEMS Microbiol. Lett. 189:171-175(2000).
RN [2]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=ATCC15003; PubMed=11048953; DOI=10.1016/S1074-5521(00)00011-9;
RX MEDLINE=20500448; Chen M., Edwards D.J., Shen B.;
RA Du L., Sanchez C., Chen M., Edwards D.J., Shen B.;
RT "The biosynthetic gene cluster for the antitumor drug bleomycin from
RT Streptomyces verticillus ATCC15003 supporting functional interactions
RT between nonribosomal peptide synthetases and a polyketide synthase.";
RL Chem. Biol. 7:623-642(2000).
CC -!- SIMILARITY: Contains 1 HTH arsr-type DNA-binding domain.
DR EMBL; AF210249; AAG02351.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001845; HTH_ArsR.
DR Pfam; PF01022; HTH_5; 1.
DR PRINTS; PR00778; HTHARSR.
DR SMART; SM00418; HTH_ARSR; 1.
KW DNA-binding; Transcription; Transcription regulation.
SQ SEQUENCE 113 AA; 12495 MW; 0D98F1C678F0437 CRC64;

Query Match 46.7%; Score 251.5; DB 2; Length 113;
Best Local Similarity 48.1%; Pred. No. 4.6e-15;
Matches 50; Conservative 26; Mismatches 27; Indels 1; Gaps 1;

QY 5 PLYKLKAEFFKTLAHPARIRILELLVERDRSVGELLSSDVGLSSNLSQLGLVLRAGVV 64
Db 4 PLYQAKAEFFKMLGHPVRIRVLELLQDGPMPVRDLIAA-IEIEPSALSQQLAVLRSGIV 62

QY 65 AARRDGNAMITYSIAAPDIAELLAVARKVLARVLSDRVAVLEDLR 108
Db 63 TSTRGTGVVYELAGGDVAELMSAARRILTEMLNGQHELLELR 106

RESULT 5
Q8FR22 PRELIMINARY; PRT; 133 AA.
AC Q8FR22;
DT 01-WAR-2003 (TrEMBLrel. 23, Created)
DT 01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative transcription regulator.
GN OrderedLocusNames=CE0945;
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314;
RX MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;
RA Nishio Y., Nakamura Y., Kawarabayasi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens.";
RL Genome Res. 13:1572-1579(2003).
CC -!- SIMILARITY: Contains 1 HTH arsr-type DNA-binding domain.
DR EMBL; AP005217; BAC17755.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001845; HTH_ArsR.
DR Pfam; PF01022; HTH_5; 1.
DR PRINTS; PR00778; HTHARSR.
DR SMART; SM00418; HTH_ARSR; 1.
KW Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 133 AA; 14636 MW; E56EF1248AF525B9 CRC64;

Query Match 40.8%; Score 220; DB 2; Length 133;
Best Local Similarity 45.0%; Pred. No. 3.7e-12;

```

```

Matches 50; Conservative 21; Mismatches 34; Indels 6; Gaps 2;

QY 5 PLYKLKAEFFKTLAHPARIRILELLVERDRSVGELLSSDVGLSSNLSQLGLVLRAGVV 64
Db 8 PLYERKANLFKGLAHPYRIRILEILSTESQVPSAMIQETGLSSHLSQHLAVLRKGLV 67

QY 65 AARRDGNAMITYSIAAPDIAELLAVARKVLARVLSDRVAVLEDLR 109
Db 68 TERNANAVSYSLTHPOVADLLRTARALNQMLAHSSDOLTSISTLPDIEA 118

RESULT 6
Q8NS13 PRELIMINARY; PRT; 104 AA.
AC Q8NS13; Q6M6S2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Predicted transcriptional regulators (Bacterial regulatory protein,
DE arsr fam).
GN OrderedLocusNames=Cgl0871, cg0993;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025 Nakagawa;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;
RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
RA Burkovski A., Dusch N., Eggelein L., Eikmanns B.J., Gaigalat L.,
RA Goesmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puhler A.,
RA Rey D.A., Rueckert C., Rupp O., Sahm H., Wendisch V.F., Wiegand I.,
RA Tauch A.;
RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
RT and its impact on the production of L-aspartate-derived amino acids
RT and vitamins.";
RL J. Biotechnol. 104:5-25(2003).
CC -!- SIMILARITY: Contains 1 HTH arsr-type DNA-binding domain.
DR EMBL; AP005276; BAB98264.1; -.
DR EMBL; EX927150; CAF19576.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001845; HTH_ArsR.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF01022; HTH_5; 1.
DR PRINTS; PR00778; HTHARSR.
DR SMART; SM00418; HTH_ARSR; 1.
KW Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 104 AA; 11850 MW; COAFA32E2AA67E50 CRC64;

Query Match 36.0%; Score 194; DB 2; Length 104;
Best Local Similarity 40.6%; Pred. No. 6.4e-10;
Matches 39; Conservative 29; Mismatches 26; Indels 2; Gaps 2;

QY 5 PLYKLKAEFFKTLAHPARIRILELLVERDRSVGELLSSDVGLSSNLSQLGLVLRAGVV 63
Db 8 PIYQKADLFKGLAHPYRIRALEIATNNEVSGQ-ITKEMDLEASHVSQHLVKRKFGL 66

QY 64 VAARDGNAMITYSIAAPDIAELLAVARKVLARVLS 99
Db 67 VSSQREGLVVYRLTYPEVADFVKVRSLLKRWAGE 102

```

Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
 -!- SIMILARITY: Contains 1 HTH arsr-type DNA-binding domain.
 DR EMBL: AP006840; BAD39395.1; -.
 DR InterPro: IPR001845; HTH_ArsR.
 DR InterPro: IPR009058; Wing_hlx_DNA_bnd.
 DR Pfam: PF01022; HTH_5; 1.
 DR PRINTS: PR00778; HTHARSR.
 DR SMART: SM00418; HTH_ArsR; 1.
 KW DNA-binding; Transcription; Transcription regulation.
 SQ SEQUENCE 99 AA; 11247 MW; 0BC610933504AE5B CRC64;

Query Match 27.9%; Score 150.5; DB 2; Length 99;
 Best Local Similarity 35.0%; Pred. No. 5.1e-06;
 Matches 35; Conservative 25; Mismatches 39; Indels 1; Gaps 1;

QY 9 LKAEFFKTLAHPARIRILELLEVERDRSVGELSSDVGLSSNLSQQLGVLRRAGVVAARR 68
 Db 1 MGARLCQVLASPKRLLELYTKDQEMTAGELAKA-VDTWPNLSQHLSEMKQGLVLSRK 59
 QY 69 DGNAMIYSIAAPDIAELAVARKVLARVLSRVAVLEDLR 108
 Db 60 EGLNMYRVASHQILETLASVRKVLVEQLAKQSLLEQVQ 99

RESULT 9
 ID Q98J11 PRELIMINARY; PRT: 112 AA.
 AC Q98J11;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Transcriptional regulator.
 GN OrderedLocusNames=mlr2159;
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Susamoto S.,
 RA Watanabe A., Idegawa K., Ishikawa K., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohata M., Matsumoto M., Matsumoto A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.,
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 CC -!- SIMILARITY: Contains 1 HTH arsr-type DNA-binding domain.
 DR EMBL: AP002999; BAB49355.1; -.
 DR GO: GO:0005622; C:intracellular; IEA.
 DR GO: GO:0003700; F:transcription factor activity; IEA.
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro: IPR001845; HTH_ArsR.
 DR InterPro: IPR009058; Wing_hlx_DNA_bnd.
 DR Pfam: PF01022; HTH_5; 1.
 DR PRINTS: PR00778; HTHARSR.
 DR SMART: SM00418; HTH_ArsR; 1.
 KW Complete proteome; DNA-binding; Transcription;
 KW Transcription regulation.
 SQ SEQUENCE 112 AA; 12422 MW; C342CB4B4A256E04 CRC64;

Query Match 27.2%; Score 146.5; DB 2; Length 112;
 Best Local Similarity 44.2%; Pred. No. 1.3e-05;
 Matches 34; Conservative 12; Mismatches 30; Indels 1; Gaps 1;

QY 11 ABEFFKTLAHPARIRILELLEVERDRSVGELSSDVGLSSNLSQQLGVLRRAGVVAARRDG 70
 Db 25 AALLKTLAHPRLMLVCTLVGEYSVSO-LEETLDIHPNLSQLTLVRDANIVETRDG 83
 QY 71 NAMIYSIAAPDIAELLA 87


```
Db      84 KQIFYRLTAEKAAQLVA 100
RESULT 10
Q8FWJ7 PRELIMINARY; PRT; 107 AA.
AC Q8FWJ7
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Transcriptional regulator, Arsr family.
GN OrderedLocusNames=BR04056;
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; PubMed=12271122; DOI=10.1073/pnas.192319099;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.B., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.P., Madupu R.,
RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J.A.,
RA Van Aken S.E., Riedmuller S., Tettelin H., Gill S.R., White O.,
RA Salzberg S.B., Hoover D.L., Lindler L.E., Hailing S.M., Boyle S.M.,
RA Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13148-13153(2002).
CC !- SIMILARITY: Contains 1 HTH arsr-type DNA-binding domain.
DR EMBL; AB014544; AAN33650.1; -.
DR TIGR; BR04056; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001845; HTH_ArsR.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF01022; HTH 5; 1.
DR PRINTS; PR00778; HTHARSR.
DR SMART; SM00418; HTH ARSR; 1.
DR Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 107 AA; 12119 MW; BEB782C953581841 CRC64;

Query Match 27.0%; Score 145.5; DB 2; Length 107;
Best Local Similarity 43.4%; Pred. No. 1.6e-05;
Matches 33; Conservative 12; Mismatches 30; Indels 1; Gaps 1;

QY 11 AEFKTLAHPARIRILELVERDRSVGELLSSDVGLESSLNLSQQLGVLRRAGVVAARRDG 70
Db 19 AELLKTLSPARLMLACTLAEGYSVGE-LEEMLGIRQPTLSQQLGVLREADIVETRREA 77

QY 71 NAMIYSIAAPDIAELL 86
Db 78 KQIYYRLVETKAAKLI 93

RESULT 11
Q8YBS6 PRELIMINARY; PRT; 117 AA.
AC Q8YBS6
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE TRANSCRIPTIONAL REGULATOR, ARSR FAMILY.
GN OrderedLocusNames=BME110810;
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688; DOI=10.1073/pnas.221575398;
RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mujer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyripides N.C., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
CC !- SIMILARITY: Contains 1 HTH arsr-type DNA-binding domain.
DR EMBL; AS003714; AAL54052.1; -.
DR FIR; AI3610; AI3610.
DR HSSP; P30340; LSMT.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR Pfam; PF01022; HTH 5; 1.
DR PRINTS; PR00778; HTHARSR.
DR SMART; SM00418; HTH ARSR; 1.
KW Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 117 AA; 13106 MW; ED8D59F23B1676A1 CRC64;

Query Match 27.0%; Score 145.5; DB 2; Length 117;
Best Local Similarity 43.4%; Pred. No. 1.7e-05;
Matches 33; Conservative 12; Mismatches 30; Indels 1; Gaps 1;

QY 11 AEFKTLAHPARIRILELVERDRSVGELLSSDVGLESSLNLSQQLGVLRRAGVVAARRDG 70
Db 29 AELLKTLSPARLMLACTLAEGYSVGE-LEEMLGIRQPTLSQQLGVLREADIVETRREA 87

QY 71 NAMIYSIAAPDIAELL 86
Db 88 KQIYYRLVETKAAKLI 103

RESULT 12
Q9LCV3 PRELIMINARY; PRT; 149 AA.
AC Q9LCV3
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Paracoccus denitrificans.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Paracoccus.
OX NCBI_TaxID=266;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB17;
RX MEDLINE=95014061; PubMed=7928987;
RA Wodara C., Kostka S., Egert M., Kelly D.P., Friedrich C.;
RT "Identification and sequence analysis of soxB gene essential for
RT sulfur oxidation ability of Paracoccus denitrificans GB17.";
RL J. Bacteriol. 176:6188-6191(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=GB17;
RX MEDLINE=97405897; PubMed=9260941;
RA Wodara C., Bardischewsky F., Friedrich C.G.;
RT "Cloning and characterization of sulfite dehydrogenase, two c-type
RT cytochromes, and a flavoprotein of Paracoccus denitrificans GB17:
RT Essential role of of sulfite dehydrogenase in lithotrophic sulfur
RT oxidation.";
RL J. Bacteriol. 179:5014-5023(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=GB17;
RA Friedrich C.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
```

```

RP SEQUENCE FROM N.A.
RC STRAIN=GB17;
RA Friedrich C.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=GB17;
RA Friedrich C.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=GB17;
RA Friedrich C.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 HTH arsr-type DNA-binding domain.
DR EMBL; X79242; CAB94376.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001845; HTH_ArsR.
DR Pfam; PF01022; HTH_5; 1.
DR PRINTS; PR00778; HTHARSR.
DR SMART; SM00418; HTH_ArsR; 1.
KW DNA-binding; Hypothetical protein; Transcription;
Transcription regulation.
SQ SEQUENCE 149 AA; 16471 MW; 3A682B67B761709B CRC64;

Query Match 27.0%; Score 145.5; DB 2; Length 149;
Best Local Similarity 44.6%; Pred. No. 2.2e-05;
Matches 37; Conservative 12; Mismatches 27; Indels 7; Gaps 2;

QY 13 FPKTLAHPARIRILELLVERDSVGE---LSSDVGLESSNLSQLGLVLRAGVVAARD 69
DB 66 FLKALCHDGLRMILCYSGPKSVTELENLLSS----RQAVVSQLARLRHEGLVSARRD 121
QY 70 GNAMIYSTIAAPDIAELLAVARKV 92
DB 122 GQTIFYSLDPKVDLLAVLKL 144

RESULT 13
Q9KI48 PRELIMINARY; PRT; 91 AA.
AC Q9KI48;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hypothetical protein.
OS Aphanothece halophytica.
OC Bacteria; Cyanobacteria; Chroococcales; Aphanothece.
OX NCBI_TaxID=72020;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20312794; PubMed=10852862;
RX DOI=10.1128/JB.182.12.3336-3344.2000;
RA Bronstein M., Schutz M., Hauska G., Fadan E., Shahak Y.;
RT "Cyanobacterial sulfide-quinone reductase: cloning and heterologous expression.";
RT J. Bacteriol. 182:3336-3344 (2000).
RC -!- SIMILARITY: Contains 1 HTH arsr-type DNA-binding domain.
DR EMBL; AF242371; AAF72965.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001845; HTH_ArsR.
DR Pfam; PF01022; HTH_5; 1.
DR PRINTS; PR00778; HTHARSR.
DR SMART; SM00418; HTH_ArsR; 1.
KW DNA-binding; Hypothetical protein; Transcription;
Transcription regulation.
SQ SEQUENCE 91 AA; 10467 MW; D9741B2C6F0168DC CRC64;

```

```

Query Match 26.8%; Score 144.5; DB 2; Length 91;
Best Local Similarity 39.2%; Pred. No. 1.6e-05;
Matches 31; Conservative 17; Mismatches 30; Indels 1; Gaps 1;

QY 7 YKLAAEFKTLAHPARIRILELLVERDSVGELESSNLSQLGLVLRAGVVA 66
DB 10 YEFLASRFKLLSEPTRIILEVICGEERKVSIDICER-TGLOQANVSKQLQLRLTAGVAC 68
QY 67 RRDGNAMIYSTIAAPDIAEL 85
DB 69 RRVGTCRYRVTDQELLEKL 87

RESULT 14
Q87AD5 PRELIMINARY; PRT; 114 AA.
ID Q87AD5;
AC Q87AD5;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Transcriptional regulator Arsr family.
GN Name=hlyU; OrderedLocusNames=PD1891;
OS Xylella fastidiosa (strain Temecul / ATCC 700964).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
Xanthomonadaceae; Xylella.
OX NCBI_TaxID=183190;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22421331; PubMed=12533478;
RX DOI=10.1128/JB.185.3.1018-1026.2003;
RA Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B.,
RA Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Moon D.H.,
RA Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.R.,
RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorfi H., Tsai S.M.,
RA Carer H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,
RA Coutinho L.L., Kimura E.T., Porro E.S., Harakava R., Kuramae E.B.,
RA Marino C.L., Gigliotti E., Abreu I.L., Alves L.M.C., do Amaral A.M.,
RA Baia G.S., Bianco S.R., Brito M.S., Cannavan F.S., Celestino A.V.,
RA da Cunha A.F., Fenille R.C., Ferro J.A., Formighieri E.F., Kishi L.T.,
RA Leoni S.G., Oliveira A.R., Roia V.E. Jr., Sasaki F.T., Sena J.A.D.,
RA de Souza A.A., Truffi D., Tsukumo F., Yanai G.M., Zatos L.G.,
RA Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Scubal J.C.,
RA Kitajima J.P.;
RT "Comparative analyses of the complete genome sequences of Pierce's
RT disease and citrus variegated chlorosis strains of Xylella
RT fastidiosa.";
RL J. Bacteriol. 185:1018-1026 (2003).
CC -!- SIMILARITY: Contains 1 HTH arsr-type DNA-binding domain.
DR EMBL; AE012560; AAO29722.1; -.
DR HSRP; P30340; ISMT.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001845; HTH_ArsR.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF01022; HTH_5; 1.
DR PRINTS; PR00778; HTHARSR.
DR SMART; SM00418; HTH_ArsR; 1.
KW Complete proteome; DNA-binding; Transcription;
Transcription regulation.
SQ SEQUENCE 114 AA; 12961 MW; 09228BA35BC01E29 CRC64;

Query Match 26.8%; Score 144.5; DB 2; Length 114;
Best Local Similarity 42.1%; Pred. No. 2e-05;
Matches 32; Conservative 14; Mismatches 29; Indels 1; Gaps 1;

QY 11 AEFKTLAHPARIRILELLVERDSVGELESSNLSQLGLVLRAGVVAARDG 70
DB 26 ANLLKTLHPVRLMLVCTLVEGEFSGVE-LEQQIGIQGTUSQQGLVRESGIVTRRN 84
QY 71 NAMIYSTIAAPDIAELL 86

```

Db 85 KQIFYRLTEAKAAQLV 100

Search completed: October 15, 2005, 04:35:49
Job time : 26.5353 secs

Db 85 KQIFYRLTEAKAAQLV 100

RESULT 15

Q9PFB1 PRELIMINARY; PRT; 114 AA.

AC Q9PFB1

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)

DE Transcriptional regulator (ArsR family).

GN OrderedLocusNames=Xf0767;

OS Xylella fastidiosa.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;

OC Xanthomonadaceae; Xylella.

OX NCBI_TaxID=2371;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=9a5C;

RX MEDLINE=20365717; PubMed=10910347; DOI=10.1038/35018003;

RA Simpson A.J.G., Rainach F.C., Arruda P., Abreu F.A., Acencio M.,

RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,

RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,

RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,

RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,

RA Coutinho L.H., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,

RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,

RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,

RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,

RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,

RA Krieger J.E., Kuranee E.E., Laigret F., Lambais M.R., Leite L.C.C.,

RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,

RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,

RA Marques M.V., Martins A.M., Martins E.M.F., Matsukuma A.Y.,

RA Menck C.F.M., Miracca E.C., Miyaki C.V., Monteiro-Vitorello C.B.,

RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,

RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,

RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,

RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,

RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,

RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,

RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,

RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,

RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,

RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,

RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;

RT "The genome sequence of the plant pathogen Xylella fastidiosa.";

RL Nature 406:151-159(2000).

CC -!- SIMILARITY: Contains 1 HTH arsr-type DNA-binding domain.

DR EMBL; AB003917; AAF83577.1; -.

DR PIR; D82766; D82766.

DR HSSP; P30340; 1SMT.

DR GO; GO:0005622; C:intracellular; IEA.

DR GO; GO:0003700; F:transcription factor activity; IEA.

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

DR InterPro; IPR01845; HTH_ArsR.

DR InterPro; IPR009058; wing_hlx_DNA_bnd.

DR Pfam; PF01022; HTH 5; 1.

DR PRINTS; PR00778; HTHARSR.

DR SMART; SM00418; HTH_ARSR; 1.

KW Complete proteome; DNA-binding; Transcription;

KW Transcription regulation.

SW SEQUENCE 114 AA; 12989 MW; 3AB9B2C6C5A58CF7 CRC64;

Query Match 26.8%; Score 144.5; DB 2; Length 114;

Best Local Similarity 42.1%; Pred.No. 2e-05;

Matches 32; Conservative 14; Mismatches 29; Indels 1; Gaps 1;

QY 11 AEFKTLAHPARIRILELIVDRSVGLSSDVLSSNLSQGLVLRRAAGVVAARDG 70

Db 26 ANLLKTLSPVRLMLVCTLVGEFSGVE-LEQIGIGQPTLSQGLVLRSGIVETRNI 84

QY 71 NAMIYSIAAPDIAELL 86

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:20:35 ; Search time 91.45 Seconds
(without alignments)
1454.845 Million cell updates/sec

Title: US-10-617-038-4
Perfect score: 1769
Sequence: 1 MPIATPEVVAELMGQAQKNS.....SORVVQACNDLHCAGKSLTH 344

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1769	100.0	344	4	AAG81114	Mycobacterium tuberculosis
2	1769	100.0	344	6	ABU34641	Protein e
3	1769	100.0	344	6	ABU36448	Protein e
4	1769	100.0	344	8	ADI37283	M. tuberc
5	1572	88.9	345	6	ABU35760	Protein e
6	1347	76.1	344	4	AAB79272	Corynebact
7	1347	76.1	344	4	AAG92789	C. glutami
8	1332	75.3	344	6	ABU26122	Protein e
9	1234.5	69.8	340	8	ADN25646	Bacterial
10	1188.5	67.2	345	4	AAU64682	Propionib
11	1188.5	67.2	345	6	ABM61201	Propionib
12	1101	62.2	355	5	ABP65587	Bifidobac
13	679.5	38.4	359	4	AAU35447	Haemophil
14	679.5	38.4	359	6	ABU30276	Protein e
15	661	37.4	358	6	ABU49134	Protein e
16	650	36.7	359	6	ABM70232	Phototrab
17	641	36.2	358	8	ADN42818	Bacterial
18	641	36.2	358	8	ADN17846	Bacterial
19	639.5	36.2	379	7	ADF05932	Bacterial
20	636	36.0	359	6	ABU19214	Protein e
21	635.5	35.9	354	6	ABU26448	Protein e
22	635.5	35.9	360	6	ABJ25863	Aspergill
23	635.5	35.9	360	6	ABJ26463	Aspergill
24	623.5	35.2	401	7	ABO65404	Klebsiell
25	609.5	34.5	359	2	AAW96159	Fructose-

26	609.5	34.5	359	4	AAB67587	Aab67587 Amino aci
27	609.5	34.5	359	5	ABG70629	Abg70629 E. coli f
28	609.5	34.5	359	8	ADS45115	Ads45115 Bacterial
29	609	34.4	358	5	ABB09164	Abb09164 Methylomo
30	609	34.4	358	5	ABG61559	Abg61559 High grow
31	601.5	34.0	359	5	AAG80302	Ag80302 C. albica
32	601.5	34.0	359	5	ABG93331	Abg93331 C. albica
33	601.5	34.0	359	5	ABP73395	Abp73395 Candida a
34	590.5	33.4	347	8	ADS21240	Ads21240 Bacterial
35	575	32.5	358	7	ADB77524	Adb77524 U. maydis
36	571.5	32.3	358	8	ADN19793	Adn19793 Bacterial
37	551	31.1	286	2	AAI19836	Aai19836 B. burgdo
38	537	30.4	270	2	AAI19837	Aai19837 B. burgdo
39	531	30.0	359	5	ABG93141	Abg93141 S. cerevi
40	531	30.0	359	6	ABR52869	AbR52869 Protein s
41	531	30.0	359	7	ADK62344	Adk62344 Disease t
42	531	30.0	359	8	ADS43824	Ads43824 Bacterial
43	526.5	29.8	321	7	ADB77522	Adb77522 U. maydis
44	506	28.6	439	8	ADN21558	Adn21558 Bacterial
45	410	23.2	118	5	ABP34371	Abp34371 Human ORF

ALIGNMENTS

RESULT 1
AAG81114
ID AAG81114 standard; protein; 344 AA.
XX
AC AAG81114;
XX
DT 04-SEP-2001 (first entry)
XX
DE Mycobacterium tuberculosis potential drug target protein SEQ ID 165.
XX
KW Drug target; growth; organism viability; characterisation.
XX
OS Mycobacterium tuberculosis.
XX
PN WO200135317-A1.
XX
PD 17-MAY-2001.
XX
PF 13-NOV-2000; 2000WO-US031152.
XX
PR 12-NOV-1999; 99US-0165086P.
PR 12-NOV-1999; 99US-0165124P.
PR 01-FEB-2000; 2000US-0179531P.
XX
(REGC) UNIV CALIFORNIA.
Eisenberg D, Rotstein SH, Marcotte EM;
WPI; 2001-329193/34.
N-PSDB; AAH51965.

Identifying nucleotide or polypeptide sequence for use as drug target, involves providing algorithm that analyzes a functional relationship between nucleotide or polypeptide sequences, and comparing the sequences.
Disclosure; Page 158; 207pp; English.
This invention relates to a method for identifying a nucleotide or polypeptide sequence that may be a drug target, or essential for growth or viability of an organism. Polynucleotide sequences AAH51947 - AAH52092 represent DNA encoding proteins AAG81096 - AAG81241, Mycobacterium tuberculosis proteins which are potential drug targets. The DNA and protein sequences are used to illustrate the method of the invention. The method involves providing an unknown nucleotide or polypeptide sequences, and comparing it to a number of sequences along with at least one algorithm capable of analysing a functional relationship between nucleotide and polypeptide sequences. The method is useful for characterising the function of nucleic acids and polypeptides that may be

CC useful as a target for a drug or essential for the growth or viability of
CC an organism

XX Sequence 344 AA;

Query Match 100.0%; Score 1769; DB 4; Length 344;
Best Local Similarity 100.0%; Pred. No. 1e-167;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPIATPEVYAEMLGQAKONSYAFPAINCTSSSETVNAAIKGFADAGSDGIIQFSTGGAEFG 60
DB 1 MPIATPEVYAEMLGQAKONSYAFPAINCTSSSETVNAAIKGFADAGSDGIIQFSTGGAEFG 60

QY 61 SGLGVKDMVTGAVALAEFTHVIAAKYPVNVALHTDHCPCDKLDSVVRPLLALSAQRVSKG 120
DB 61 SGLGVKDMVTGAVALAEFTHVIAAKYPVNVALHTDHCPCDKLDSVVRPLLALSAQRVSKG 120

QY 121 GNPLFQSHMWDGSAVPIDENLAIQAELLKAAAAAKIILEIEIGVVGGEEDGVANEINEKL 180
DB 121 GNPLFQSHMWDGSAVPIDENLAIQAELLKAAAAAKIILEIEIGVVGGEEDGVANEINEKL 180

QY 181 YTSPEDEFEKTIIEALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDIQAQGOQVAAAKLGL 240
DB 181 YTSPEDEFEKTIIEALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDIQAQGOQVAAAKLGL 240

QY 241 PADAKPPDFVPHGGSGSLKSEIEEALRYGVVKNMVDTDTOYAFTRPIAGHMTFNYDGVK 300
DB 241 PADAKPPDFVPHGGSGSLKSEIEEALRYGVVKNMVDTDTDTOYAFTRPIAGHMTFNYDGVK 300

QY 301 VDGEVGVKKVYDPRSYLKKAEASMSQRVQACNDLHCAGKSLTH 344
DB 301 VDGEVGVKKVYDPRSYLKKAEASMSQRVQACNDLHCAGKSLTH 344

RESULT 2

ABU34641
ID ABU34641 standard; protein; 344 AA.
XX AC ABU34641;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #20168.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Mycobacterium bovis.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI; 2003-029926/02.
XX DR N-PSDB; ACA38511.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 25; SEQ ID NO 62565; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pcr_sequences

XX SQ Sequence 344 AA;

Query Match 100.0%; Score 1769; DB 6; Length 344;
Best Local Similarity 100.0%; Pred. No. 1e-167;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPIATPEVYAEMLGQAKONSYAFPAINCTSSSETVNAAIKGFADAGSDGIIQFSTGGAEFG 60
DB 1 MPIATPEVYAEMLGQAKONSYAFPAINCTSSSETVNAAIKGFADAGSDGIIQFSTGGAEFG 60

QY 61 SGLGVKDMVTGAVALAEFTHVIAAKYPVNVALHTDHCPCDKLDSVVRPLLALSAQRVSKG 120
DB 61 SGLGVKDMVTGAVALAEFTHVIAAKYPVNVALHTDHCPCDKLDSVVRPLLALSAQRVSKG 120

QY 121 GNPLFQSHMWDGSAVPIDENLAIQAELLKAAAAAKIILEIEIGVVGGEEDGVANEINEKL 180
DB 121 GNPLFQSHMWDGSAVPIDENLAIQAELLKAAAAAKIILEIEIGVVGGEEDGVANEINEKL 180

QY 181 YTSPEDEFEKTIIEALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDIQAQGOQVAAAKLGL 240
DB 181 YTSPEDEFEKTIIEALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDIQAQGOQVAAAKLGL 240

QY 241 PADAKPPDFVPHGGSGSLKSEIEEALRYGVVKNMVDTDTOYAFTRPIAGHMTFNYDGVK 300
DB 241 PADAKPPDFVPHGGSGSLKSEIEEALRYGVVKNMVDTDTDTOYAFTRPIAGHMTFNYDGVK 300

QY 301 VDGEVGVKKVYDPRSYLKKAEASMSQRVQACNDLHCAGKSLTH 344
DB 301 VDGEVGVKKVYDPRSYLKKAEASMSQRVQACNDLHCAGKSLTH 344

RESULT 3

ABU36448
ID ABU36448 standard; protein; 344 AA.
XX AC ABU36448;
XX XX
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #21975.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KW Mycobacterium tuberculosis.
 OS WO20027183-A2.
 PN 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US0009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Hasebeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI: 2003-029926/02.
 DR N-PSDB; ACA0318.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 64372; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 344 AA;
 Query Match 100.0%; Score 1769; DB 6; Length 344;
 Best Local Similarity 100.0%; Pred. No. 1e-167;
 Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 MPIATPEVYAEMLGQAKNSYAFPAINCTSSSETVNAAIKGFADAGSDGIIOFTSTGGAEGF 60
 1 MPIATPEVYAEMLGQAKNSYAFPAINCTSSSETVNAAIKGFADAGSDGIIOFTSTGGAEGF 60

QY 61 SGLGVKDMVTGAVALAEFTHTVIAAKYPNVNVALHTDHCPCPKDLSYVRPLLAISAQRVSKG 120
 DB 61 SGLGVKDMVTGAVALAEFTHTVIAAKYPNVNVALHTDHCPCPKDLSYVRPLLAISAQRVSKG 120
 QY 121 GNPLFQSHMWDGSAVPIDENLAIQELLKAAAAAKIILIEIGVVGEGEDGVANEINEKL 180
 DB 121 GNPLFQSHMWDGSAVPIDENLAIQELLKAAAAAKIILIEIGVVGEGEDGVANEINEKL 180
 QY 181 YTSPEDEFKTIETALGAGEHGKYLAAATFCGNVHGVPKGNVKLRPDI LAOQOQVAAAKLGL 240
 DB 181 YTSPEDEFKTIETALGAGEHGKYLAAATFCGNVHGVPKGNVKLRPDI LAOQOQVAAAKLGL 240
 QY 241 PADAKPDEFVHFGSGSLKSEIEBALRYGVVKNVDTDTQYAFTRPIAGHMTFTNYDGVLK 300
 DB 241 PADAKPDEFVHFGSGSLKSEIEBALRYGVVKNVDTDTQYAFTRPIAGHMTFTNYDGVLK 300
 QY 301 VDGEVGVKKVYDPRSYLKKAEASMSQVVQACNDLHCAGKSLTH 344
 DB 301 VDGEVGVKKVYDPRSYLKKAEASMSQVVQACNDLHCAGKSLTH 344

RESULT 4
 ADI37283
 ID ADI37283 standard; protein; 344 AA.
 XX
 AC ADI37283;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE M. tuberculosis low oxygen induced antigen RV0363c SEQ ID NO:4.
 XX
 KW mycobacterial infection; vaccine; tuberculosis;
 KW Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;
 KW Low oxygen induced antigen.
 XX
 OS Mycobacterium tuberculosis.
 PN WO2004006952-A2.
 XX
 PD 22-JAN-2004.
 XX
 PF 08-JUL-2003; 2003WO-DK000477.
 XX
 PR 13-JUL-2002; 2002DK-00001099.
 XX
 PA (STAT-) STATENS SERUM INST.
 XX
 PI Andersen P, Rosenkrands I, Stryhn A;
 XX
 DR WPI: 2004-122778/12.
 DR N-PSDB; ADI37328.
 XX
 PT Use of one or more polypeptides or their fragments, which are expressed
 PT during the latent stage of the mycobacterial infection, and/or nucleic
 PT acids encoding the polypeptides, for a therapeutic vaccine against
 PT tuberculosis.
 XX
 PS Claim 3; SEQ ID NO 4; 76pp; English.
 XX
 CC The present invention describes polypeptides or their fragments, which
 CC are expressed during the latent stage of a mycobacterial infection,
 CC and/or nucleic acids encoding the polypeptides, which are useful for
 CC creating a therapeutic vaccine against tuberculosis. Also described: (1)
 CC a therapeutic vaccine against tuberculosis comprising one or more
 CC polypeptides; (2) a method for treating an animal, including a human
 CC being, with tuberculosis caused by virulent mycobacteria, e.g. by
 CC Mycobacterium tuberculosis, *M. africanum* or *M. bovis*; (3) a method for
 CC immunising an animal, including a human being, against tuberculosis
 CC caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis
 CC caused by virulent mycobacteria in an animal, including a human being;
 CC (5) a method for diagnosing previous or ongoing infection with a virulent
 CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
 CC infection in a subject. The polypeptides have antibacterial activities,

CC and can be used in vaccines and in gene therapy. The polypeptides are
 CC useful for the manufacture of a therapeutic vaccine for treating an
 CC individual who is infected by a virulent mycobacterium, e.g. M.
 CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
 CC The present sequence represents a low oxygen induced antigen, which is
 CC used in the exemplification of the present invention.

XX Sequence 344 AA;

Query Match 100.0%; Score 1769; DB 8; Length 344;
 Best Local Similarity 100.0%; Pred. No. 1e-167;
 Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPIATPEVVAEMLGQAKONSYAFPAINCTSSTVNAAIKGFADAGSDGLIIFSTGGAFFG 60
 DB 1 MPIATPEVVAEMLGQAKONSYAFPAINCTSSTVNAAIKGFADAGSDGLIIFSTGGAFFG 60

QY 61 SGLGVKDMVTGAVALAEFTHVIAAKYPVNVVALHTDHCPCDKLDSYVRPLLAISAQORVSKG 120
 DB 61 SGLGVKDMVTGAVALAEFTHVIAAKYPVNVVALHTDHCPCDKLDSYVRPLLAISAQORVSKG 120

QY 121 GNPLFQSHMWDGSAVPIDENLAIQAELLKAAAAAKIILEIEIGVVGSGEDGVANEINEKL 180
 DB 121 GNPLFQSHMWDGSAVPIDENLAIQAELLKAAAAAKIILEIEIGVVGSGEDGVANEINEKL 180

QY 181 YTSPEDEKTIIEALGAGEHGKYLAAATFGNVHGVKPGNVKLRPDLAQGGQVAAAKLGL 240
 DB 181 YTSPEDEKTIIEALGAGEHGKYLAAATFGNVHGVKPGNVKLRPDLAQGGQVAAAKLGL 240

QY 241 PADAKPFDVFPHGGSGSLKSEIEEALRYGVVWKMVNDTDTQYAFTRPIAGHMFTNYDGVLYK 300
 DB 241 PADAKPFDVFPHGGSGSLKSEIEEALRYGVVWKMVNDTDTQYAFTRPIAGHMFTNYDGVLYK 300

QY 301 VDGEVGVKKVYDPRSYLKAESMSQVQVQACNDLHCAGKSLTH 344
 DB 301 VDGEVGVKKVYDPRSYLKAESMSQVQVQACNDLHCAGKSLTH 344

RESULT 5

ABU35760
 ID ABU35760 standard; protein; 345 AA.

XX AC ABU35760;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #21287.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Mycobacterium leprae.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haseibeck R, Ohlsen KL, Zyskind JW;
 XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI: 2003-029226/02.

XX DR N-PSDB; ACA39630.

XX PT New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 63684; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 345 AA;

Query Match 88.9%; Score 1572; DB 6; Length 345;
 Best Local Similarity 87.7%; Pred. No. 5.1e-148;

Matches 300; Conservative 25; Mismatches 17; Indels 0; Gaps 0;

QY 1 MPIATPEVVAEMLGQAKONSYAFPAINCTSSTVNAAIKGFADAGSDGLIIFSTGGAFFG 60

DB 1 MPIATPEVVAEMLRRAKENSYAFPAINCTSSTVNAAIKGFADAGSDGLIIFSTGGAFFG 60

QY 61 SGLGVKDMVTGAVALAEFTHVIAAKYPVNVVALHTDHCPCDKLDSYVRPLLAISAQORVSKG 120

DB 61 SGLGVKDMVTGAVALAEFTHTTAAKYPINVALHTDHCPCDKLDSYVRPLLAISARRVATG 120

QY 121 GNPLFQSHMWDGSAVPIDENLAIQAELLKAAAAAKIILEIEIGVVGSGEDGVANEINEKL 180

DB 121 KDLFGSHMWDGSAIPIDENLAIQAELLKAAAAAKIILEIEIGVVGSGEDGVANEINEKL 180

QY 181 YTSPEDEKTIIEALGAGEHGKYLAAATFGNVHGVKPGNVKLRPDLAQGGQVAAAKLGL 240

DB 181 YTPKDFVKTIDALGAGEHGKYLAAATFGNVHGVKPGNVKLRPDLAQGGQVAAAKLSQ 240

QY 241 PADAKPFDVFPHGGSGSLKSEIEEALRYGVVWKMVNDTDTQYAFTRPIAGHMFTNYDGVLYK 300

DB 241 SEGSKPFDVFPHGGSGSEKSEIEEALRYGVVWKMVNDTDTQYAFTRPVSHMFTNYDGVLYK 300

QY 301 VDGEVGVKKVYDPRSYLKAESMSQVQVQACNDLHCAGKSL 342

DB 301 VDGEVGNKKVYDPRSYLKAESMSQVQVQACNDLHCAGKSV 342

RESULT 6

AAB79272

ID AAB79272 standard; protein; 344 AA.

XX AC AAB79272;

Db 244 --KEKFDLVFRGSGSTQAEIADAVSYGVIRKMNIDTQYAFSRVVEHMFKNYDMLK 301
QY 301 VDEGVGVKKYDPRSLKKAESMSORVVOACNDLHCAGKSLT 343
Db 302 IDEGVGNKKYDPRSGKKGAGMAARIVEACEQLGSKGTSVS 344

RESULT 12
ABP65587
ID ABP65587 standard; protein; 355 AA.
XX
AC ABP65587;
XX
DT 19-NOV-2002 (first entry)
XX
DE Bifidobacterium longum NCC2705 ORF amino acid sequence SEQ ID NO:331.
XX
KW Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
KW antiarrheic; antibacterial; inhibitor of Salmonella; detection;
KW identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
KW rotavirus; food composition; pharmaceutical composition.
XX
OS Bifidobacterium longum.
XX
XX EP1227152-A1.
XX
PD 31-JUL-2002.
XX
PF 30-JAN-2001; 2001EP-00102050.
XX
PR 30-JAN-2001; 2001EP-00102050.
XX
PA (NEST) SOC PROD NESTLE SA.
XX
XX WPI; 2002-668397/72.
XX
DR Novel polynucleotide comprising Bifidobacterium genome sequence useful as
PT a probe or primer for detecting and/or identifying Bifidobacterium longum
PT in a biological sample.
XX
PS Claim 3; SEQ ID NO 331; 80pp; English.
XX

The present invention describes a polynucleotide (I) comprising a
CC sequence of a Bifidobacterium genome selected from the nucleotide
CC sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at
CC least 90% identity or which hybridises with the sequences given in
CC ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding a
CC fusion protein, comprising a sequence selected from 1097 sequences given
CC in ABP65258 to ABP66354 ligated in frame to a polynucleotide encoding a
CC heterologous polypeptide. (I) has antiarrheic and antibacterial
CC activities, and can be used as an inhibitor of Salmonella. (I) (which is
CC a probe) is useful for the detection and/or identification of
CC Bifidobacterium longum in a biological sample. A carrier containing the
CC lactic acid bacterium Bifidobacterium longum NCC2705 (CNCM I-2618) can be
CC used for preventing and/or treating diarrhoea brought about by pathogenic
CC bacteria and/or rotavirus. The carrier is a food composition selected
CC from milk, yogurt, curd, cheese, fermented milks, milk based fermented
CC products, ice-creams, fermented cereal based products, milk based
CC powders, infant formula, pet food or a pharmaceutical composition
CC selected from tablets, liquid bacterial suspensions, dried oral
CC supplement, wet oral supplement, dry tube feeding or wet tube feeding.
CC (I) is useful in DNA arrays or chips to carry out analysis of the
CC expression of the Bifidobacterium gene. ABQ81844 to ABQ81850 represent
CC Bifidobacterium related nucleotide sequences given in the Sequence
CC Listing from the present invention but not mentioned further within the
CC specification. N.B. The sequence data for this patent is not represented
CC in the printed specification but is based on sequence information
CC supplied by the European Patent Office
XX
SQ Sequence 355 AA;

Query Match

62.2%; Score 1101; DB 5; Length 355;

Best Local Similarity 61.9%; Pred. No. 6.6e-101;
Matches 219; Conservative 42; Mismatches 81; Indels 12; Gaps 2;

QY 1 MBIATPEVVAEMLGOKNSYAFPAINTCTSSSTVNAAIKGFADAGSDGIIQFSTGGAERG 60
Db 1 MTIATPERVAEMLAARRGGYAYPAINTVSSOTLNAALKGFADAGSDGIIQISVGGAAYV 60

QY 61 SGLGVKDMVTGAVALAEFTHTVIAKYP-VNVALHTDHCPOKLDYSYVRPLLAISAQRVSK 119
Db 61 SGRGVNDRVTGSLAALAAFAHEVAAYKPNITIALHTDHCQYLDKDEWARPELLAHEADQVAR 120

QY 120 GGNPLFQSHWDSAPIDENLAIQAELLKAAAIAKILIEIGVVGGEEDGVANEINEK 179
Db 121 GOEPTFQSHWDSAPIDENLAIQAELLKAAAIAKILIEIGVVGGEEDGVANEINEK 180

QY 180 LYTSPEDFEKTTIEALGAGEHGKYLAAATFGNVHGVVYKPGNVKLRPDIL-----A 228
Db 181 LYSTPEDGLEVARRLGLGERGYMAAFTGNVHGAYKPGVVKLRPSLLGDIQARVARAVA 240

QY 229 QGQVAAAKLGLPADAKPDPFVHGGSGSLKSEIBEARLYGVVXNMVDTDTQYAFTRPIA 288
Db 241 EGELPSAAGIVDPNGKPPPELVHGGSGSRPEIAEAVSYGVIKMINDTDTQYAFTRPIA 300

QY 289 GHMFTNYDGVKVDGVEGVKVVDPERSYLLKKAESMSORVVOACNDLHCAGKSL 342
Db 301 DHVFENYDKVLKIDGVEGVEKFPDPRSGRKAEDSMSARVWEACROLGSGAKAL 354

RESULT 13
AAU35447
ID AAU35447 standard; protein; 359 AA.
XX
AC AAU35447;
XX
DT 14-FEB-2002 (first entry)
XX
DE Haemophilus influenzae cellular proliferation protein #88.
XX
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW antibacterial; drug design.
XX
OS Haemophilus influenzae.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS53306.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Example 3; SEQ ID NO 11040; 51pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes

Db 247 EFVSKERNLP--AKPINFVHGSGSSREIREAIGYGAIKMNDTDTQMA----- 295
QY 292 FTNYDGLK-----VDGEVG-----VKKVVDPSPSYLKABASMSQRVVQACNDL 335
Db 296 --SWNGILNFYKANEAYLQOLNPEGPDAPNKYYDPRVWLKMBESMSKRLQSFEDL 353
QY 336 HC 337
Db 354 NC 355

RESULT 15

ABU49134
ID ABU49134 standard; protein; 358 AA.

AC

XX ABU49134;

DT 19-JUN-2003 (first entry)

DE Protein encoded by prokaryotic essential gene #34661.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Vibrio cholerae.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX N-PSDB; ACA53004.

XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 77058; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism's activity; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 358 AA;

Query Match 37.4%; Score 661; DB 6; Length 358;

Best Local Similarity 43.2%; Pred. No. 6.5e-57;

Matches 145; Conservative 59; Mismatches 110; Indels 22; Gaps 7;

QY 16 AKONSVAFFAINCTSETVNAAIKGFADAGSDGIIQFSTGGAEFGSLGVK-----DMVT 70

Db 25 AKONGFALPAVNCVGTSDSVNGVLEAAAKVKAPVIVQFSNGGAFFAGKGLKLEGQQAQIL 84

QY 71 GAVAAAEFTHVIAAKYPNVNVALHTDHCPCDKLDSYVRPLLAISAQVSKGNLFLQSHMW 130

Db 85 GAIAAGAKYVHTVAEAYGVPVILHTDHAANKLL-PRIDGLLDAGEKHFAETGKPLFSSHMI 143

QY 131 DGSAPVIDENLAIAGELLKAAAAAKIILIEIGVVGGEEGDVANE--INEKLYTSPEDFE 188

Db 144 DLSEESLEENIEISAKYLERMAKNMNTLIELGCTGGEEDVDNSDMSSELYTSPEDVA 203

QY 189 KTIEALGAGEHGKYLAAATFGNVHGVKPGNVKLRPDILAOGOOAAAKLGLPADAKPFD 248

Db 204 YAYEKLNAISH-RFTIAASFGNVHGVKPGNVVLTTLIRDSQKYVSEKFGLPENS--LN 260

QY 249 FVFHGGSGSLKSEIEALRYGVVKNVNDTDTQVAFTRPIAGHMFTNYDGVKLVGDEYV-- 306

Db 261 FVFHGGSGSSLEIREISYGVIKNNIDTDTQWATWDGIRQYEAKNHD---YLQGGQIGNP 317

QY 307 -----VKKYDPRSYLKKAEASMSQRVVQACNDLH 336

Db 318 TGEDAPNKKYYDPRVWLRAGQAAMVARLEKAFSDLN 353

Search completed: October 15, 2005, 04:26:18

Job time : 93.45 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 01:42:44 ; Search time 28.2439 Seconds
(without alignments)
909.197 Million cell updates/sec

Title: US-10-617-038-4

Perfect score: 1769

Sequence: 1 MPIATPEVYAEMLGQAKQNS.....SORVVOACNDLHCAGKSLTH 344

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/PCtUS_COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	679.5	38.4	359	4	US-09-403-343B-21	Sequence 21, Appl
2	639.5	36.2	379	4	US-09-543-681A-6217	Sequence 6217, A
3	623.5	35.2	401	4	US-09-489-039A-11921	Sequence 11921, A
4	609.5	34.5	359	4	US-09-098-219B-2	Sequence 2, Appl
5	609.5	34.5	359	4	US-10-164-204-2	Sequence 2, Appl
6	609.5	34.5	359	4	US-09-923-109-2	Sequence 2, Appl
7	609.5	34.4	358	4	US-09-934-901-8	Sequence 8, Appl
8	609.5	34.4	358	4	US-09-934-868-18	Sequence 18, Appl
9	609.5	34.4	358	4	US-10-321-210-8	Sequence 8, Appl
10	609.5	34.4	358	4	US-10-320-874-8	Sequence 8, Appl
11	602.5	34.1	366	4	US-09-248-796A-17059	Sequence 17059, A
12	327.18	29.3	293	4	US-09-583-110-5316	Sequence 5316, Ap
13	327.18	30.3	303	4	US-09-107-433-4783	Sequence 4783, Ap
14	309.17	30.3	303	3	US-09-134-001C-4256	Sequence 4256, Ap
15	304.5	17.2	281	4	US-09-710-279-1026	Sequence 1026, Ap
16	302.17	28.6	4	US-10-138-701-18	Sequence 18, Appl	
17	293.16	31.9	4	US-09-107-532A-5442	Sequence 5442, Ap	
18	266.5	15.1	287	4	US-09-543-681A-7123	Sequence 7123, Ap
19	266.5	15.1	298	4	US-09-489-039A-12846	Sequence 12846, A
20	247.14	14.0	290	4	US-09-107-532A-7033	Sequence 7033, Ap
21	244.13	13.8	290	4	US-09-107-532A-4005	Sequence 4005, Ap
22	238.5	13.5	348	4	US-09-540-236-2479	Sequence 2479, Ap
23	233.5	13.2	364	4	US-09-252-991A-19881	Sequence 19881, A
24	231.5	13.1	357	4	US-09-328-352-7833	Sequence 7833, Ap
25	231.13	13.1	286	4	US-09-711-164-385	Sequence 385, App
26	221.5	12.5	335	4	US-09-934-901-6	Sequence 6, Appl
27	221.5	12.5	335	4	US-09-934-868-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1

US-09-403-343B-21
; Sequence 21, Application US/09403343B
; Patent No. 6555091
; GENERAL INFORMATION:
; APPLICANT: JOLIVET-REYNAUD, COLETTE
; APPLICANT: PERRON, HERVE
; APPLICANT: MANDRAND, BERNARD
; TITLE OF INVENTION: POLYPEPTIDE CAPABLE OF REACTING WITH ANTIBODIES OF
; TITLE OF INVENTION: PATIENTS SUFFERING FROM MULTIPLE SCLEROSIS AND USES
; FILE REFERENCE: 104574
; CURRENT APPLICATION NUMBER: US/09/403,343B
; CURRENT FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: FR/97/05679
; PRIOR FILING DATE: 1997-04-29
; PRIOR APPLICATION NUMBER: FR/97/16870
; PRIOR FILING DATE: 1997-12-31
; PRIOR APPLICATION NUMBER: PCT/FR98/00870
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 21
; LENGTH: 359
; TYPE: PRT
; ORGANISM: MSRV
US-09-403-343B-21

Query Match 38.4%; Score 679.5; DB 4; Length 359;

Best Local Similarity 42.5%; Pred. No. 8.9e-64;

Mismatches 154; Conservative 61; Mismatches 102; Indels 45; Gaps 10;

QY 3 IATPEVYAEMLGQAKQNSYAFPAINTSETVNAAIKGFADAGSDGIQFSTGAEFGSG 62

Db 12 VVTGEDVQKVFAYAKEHNEFAVNCVSDSVNAVLETAARVKAPVIOFSGNGAAFYAG 71

QY 63 LGVKOM-----VTGAVALAEFTHTVIAAKYPVNAVLTHTCHPKKLDYSVVRPLLAISAOR 116

Db 72 KGKIPSTGTRPDVLGAIAAGAKQVHTLUKEYGVVPLIHTDRAAK-KULPPIDGLDAGEKH 130

QY 117 VSKGNPLFQSHWMDGSAPVIDENLAIAQELKAAAAAKIILIEICVWVGGEEDGVAN-E 175

Db 131 FAETGRPLFSSHMDLSEESMEENMAICREYLARMKMGMTLEIGITGTEEDGVNDSD 190

QY 176 INE-KULTSPDEFKTEIALGAGEH---GKYLAAATFGNVHGVYKPGNVKLRPDILAAQGG 231

Db 191 VDESRLYTPQSDVLYVYDQL----HPVSPNFTVAAAFGNVHGVYKPGNVKLRKPSILGESQ 246

QY 232 QVAAAKLGLPADAKPFDVPHGSGSLKSEIEEALRYGVVKNMVDTDTOVAFTRPIAGHM 291

Db 247 EFVSKERNLP--AKPINFPHGSGSSRBEIRAIYGAIKMNIDTDTOWA----- 295

```
QY 292 FTNYDGVLK-----VDGEVG-----VKVYDPRSYLKAAEASMSORVVQACNDL 335
; SEQ ID NO 11921
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11921

Db 296 --SWGILNFYKANEAYLQGLGNPEGDPAPNKKYIDPRVWLRLKMEESMSKRLQSFEDL 353
;
QY 336 HC 337
;
Db 354 NC 355
;

Query Match 35.2%; Score 623.5; DB 4; Length 401;
Best Local Similarity 41.4%; Pred. No. 1e-57;
Matches 143; Conservative 55; Mismatches 108; Indels 39; Gaps 8;

RESULT 2
US-09-543-681A-6217
; Sequence 6217, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6217
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6217
```

```
Query Match 36.2%; Score 639.5; DB 4; Length 379;
Best Local Similarity 40.8%; Pred. No. 1.8e-59;
Matches 146; Conservative 58; Mismatches 115; Indels 39; Gaps 8;

QY 3 1ATPEYAEMLGQAKONSYAFPAINCTSETVNAAIKGFADAGSDGIIQFSTGGAEFGSG 62
;
Db 32 VITGDVQKVFVAKENNFPALPANCVTGDSINAVLEAAAKVRSPIVQFSNGGAFTAG 91
;
QY 63 LGVK-----DMVTGAVALAEFTHVIAAKYPNVNVALHTDCHPKDLSYVRPLLAISQR 116
;
Db 92 KGLKAEAPQQAAILGAISGAHHVQMAEYGPVILHTDHCAC-KLLPWIIDGLLDAGEKH 150
;
QY 117 VSKGNPLFQSHMWDGSAVPIDENLAIQAELLKAAAAAKIILEIGVVGGEEDGVANE- 175
;
Db 151 YAKTKGPIFSSHMIDLSSELEENIEICSKYLARMAKMTLEIELGCTGGEEDGVNTG 210
;
QY 176 -INEKLYTSPDEFKTIKALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDLAQQQVVA 234
;
Db 211 MDSSALYTPQEDVAVAYEKLNA-ISPRFTIAASFGNVHGVYKPGNVQLTPKILRNSQYV 269
;
QY 235 AAKLGLPADAKPFDVEFHGGSSLSKSEIEEALRYGVVKNVDTDTQYAFTRPIAGHMFTN 294
;
Db 270 SEKNYLPHNS--LNFVHHGGSSAAEIEKAEVSGVGVKNMIDTDTQWA-----T 316
;
QY 295 YDGVLK-----VDGEVG-----VKVYDPRSYLKAAEASMSORVVQACNDLH 336
;
Db 317 WDGILOFYKKNEGYLSQGLNPEGADKNKKYIDPRVWLRLHCQTSMVVRLQAFKELN 374
;
```

```
RESULT 3
US-09-489-039A-11921
; Sequence 11921, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
```

```
; SEQ ID NO 11921
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11921

Query Match 35.2%; Score 623.5; DB 4; Length 401;
Best Local Similarity 41.4%; Pred. No. 1e-57;
Matches 143; Conservative 55; Mismatches 108; Indels 39; Gaps 8;

QY 16 AKONSYAPPAINCTSETVNAAIKGFADAGSDGIIQFSTGGAEFGSLGVKX-----MV 69
;
Db 67 AKENNFPALPANCVTGDSINAVLEAAAKVRSPIVQFSNGGAFTAGKVTDPVQGAAI 126
;
QY 70 TGAVALAEFTHVIAAKYPNVNVALHTDCHPKDLSYVRPLLAISQVRVSKGNPLFQSHM 129
;
Db 127 LGATSGAAHHVQMAEYGPVILHTDHCAC-KLLPWIIDGLLDAGEKHFAATGKPLFSSHM 185
;
QY 130 WDSGAVPIDENLAIQAELLKAAAAAKIILEIGVVGGEEDGVANEINE--KLYTSPEDF 187
;
Db 186 IDLSSELSHENIEICSKYLARMAKMTLEIELGCTGGEEDGVNDSHMDASALYTPQEDV 245
;
QY 188 EKTIEALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDLAQQQVAAAKLGLPADAKPF 247
;
Db 246 DYAYTEL-SKISPRFTIAASFGNVHGVYKPGNVVLTPTILRDSQEVVSKKHNLPHNS--L 302
;
QY 248 DFVHHGGSSLSKSEIEEALRYGVVKNVDTDTQYAFTRPIAGHMFTNVDGVLK----- 300
;
Db 303 NLFVHHGGSSAAEIEKAEVSGVGVKNMIDTDTQWA-----TWGILQYVYKANE 351
;
QY 301 -VDGEVG-----VKVYDPRSYLKAAEASMSORVVQACNDLH 336
;
Db 352 YLQQLGNPKGEDQPNKKYIDPRVWLRLAAQTSMTVRLQAFKELN 396
;

RESULT 4
US-09-098-219B-2
; Sequence 2, Application US/09098219B
; Patent No. 6441277
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard
; APPLICANT: Cheikh, No. 6441277dine
; APPLICANT: Kishore, Ganesh
; TITLE OF INVENTION: Expression of Fructose 1,6 Bisphosphate
; TITLE OF INVENTION: Aldolase in Transgenic Plants
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: US
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/098,219B
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/049,995
; FILING DATE: 17-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kammerer, Patricia A.
; REGISTRATION NUMBER: 29,775
; REFERENCE/DOCKET NUMBER: MOBT-086
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-787-1400
; TELEFAX: 713-787-1440
; INFORMATION FOR SEQ ID NO: 2;
```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 359 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-09-098-2198-2
Query Match 34.5%; Score 609.5; DB 4; Length 359;
Best Local Similarity 40.3%; Pred. No. 2.7e-56;
Matches 139; Conservative 58; Mismatches 109; Indels 39; Gaps 8;

QY 16 AKONSAFPAINCTSETVNAAIKGFADAGSDGIIQFSTGGAEFGSLGVKD-----MV 69
Db 25 AKENNFALPAVNCVGTDSINAVLETAAKVKAPVIVQFSNGGASFIAGKGVKSDVPQGAAI 84
QY 70 TGAVALAEETHVIAAKYPVNVALLTHDCHPKDKLDSVVRPLLATSAQRVSKGNGNPLFQSHM 129
Db 85 LGAISGAHHVHQAHEHYGVPIVILTHDCAK-KLLPWIDGLLDAGEKHFAATGKPLFSSHM 143
QY 130 WGSASVPIDENLAIAGELLKAAAAAIIIEIGVVGSGEDGVANINE--KLYTSPEDF 187
Db 144 IDLSESLQENIEICSKYLERMSKIGMTLEIELGCTGGEEDGVNDSHMDASALYTPQEDV 203
QY 188 EXTIEALGAGECHKYLLAATFGNVHGVYKPGNVKLPDPDILAQQOQVAAAKLGLPADAKPF 247
Db 204 DVAYTEL-SKISPRFTIAASFGNVHGVYKPGNVVLTPTILRDSQEVYVSKKHNLPHNS--L 260
QY 248 DFVPHGGSGSLKSEIEALRYGVVKNVDTDTQYAFTRPIAGHMFNTNYDGVLK----- 300
Db 261 NFVPHGGSGTAQEIKDSVSYGVVKNVDTDTQWA-----TWEGVLNYYKANE 309
QY 301 -VDGEVG-----VKKYDPRSYLKAEASMSQRVVOACNDLH 336
Db 310 YLQGLGNPKGEDQPNKKYDPRVLRAGQTSMIARLEKAFQELN 354

RESULT 5
US-10-164-204-2
; Sequence 2, Application US/10164204
; Patent No. 6663906
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard
; APPLICANT: Cheikh, No. 6663906dine
; APPLICANT: Kishore, Ganesh
; TITLE OF INVENTION: Expression of Fructose 1,6 Bisphosphate Aldolase in Transgenic Pl
; FILE REFERENCE: 11899.0086.DVUS02 (MOBT.086-2)
; CURRENT APPLICATION NUMBER: US/10/164,204
; CURRENT FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: 09/098,219
; PRIOR FILING DATE: 1998-06-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 359
; TYPE: PRT
; ORGANISM: E. coli
US-10-164-204-2
Query Match 34.5%; Score 609.5; DB 4; Length 359;
Best Local Similarity 40.3%; Pred. No. 2.7e-56;
Matches 139; Conservative 58; Mismatches 109; Indels 39; Gaps 8;

QY 16 AKONSAFPAINCTSETVNAAIKGFADAGSDGIIQFSTGGAEFGSLGVKD-----MV 69
Db 25 AKENNFALPAVNCVGTDSINAVLETAAKVKAPVIVQFSNGGASFIAGKGVKSDVPQGAAI 84
QY 70 TGAVALAEETHVIAAKYPVNVALLTHDCHPKDKLDSVVRPLLATSAQRVSKGNGNPLFQSHM 129
Db 85 LGAISGAHHVHQAHEHYGVPIVILTHDCAK-KLLPWIDGLLDAGEKHFAATGKPLFSSHM 143
QY 130 WGSASVPIDENLAIAGELLKAAAAAIIIEIGVVGSGEDGVANINE--KLYTSPEDF 187
Db 144 IDLSESLQENIEICSKYLERMSKIGMTLEIELGCTGGEEDGVNDSHMDASALYTPQEDV 203
```

```

QY 188 EXTIEALGAGECHKYLLAATFGNVHGVYKPGNVKLPDPDILAQQOQVAAAKLGLPADAKPF 247
Db 204 DVAYTEL-SKISPRFTIAASFGNVHGVYKPGNVVLTPTILRDSQEVYVSKKHNLPHNS--L 260
QY 248 DFVPHGGSGSLKSEIEALRYGVVKNVDTDTQYAFTRPIAGHMFNTNYDGVLK----- 300
Db 261 NFVPHGGSGTAQEIKDSVSYGVVKNVDTDTQWA-----TWEGVLNYYKANE 309
QY 301 -VDGEVG-----VKKYDPRSYLKAEASMSQRVVOACNDLH 336
Db 310 YLQGLGNPKGEDQPNKKYDPRVLRAGQTSMIARLEKAFQELN 354

RESULT 6
US-09-923-109-2
; Sequence 2, Application US/09923109
; Patent No. 6716474
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard
; APPLICANT: Cheikh, No. 6716474dine
; APPLICANT: Kishore, Ganesh
; TITLE OF INVENTION: Expression of Fructose 1,6 Bisphosphate
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Arnold White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: US
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/923,109
; FILING DATE: 06-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/098,219
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 60/049,995
; FILING DATE: 17-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kammerer, Patricia A.
; REGISTRATION NUMBER: 29,775
; REFERENCE/DOCKET NUMBER: MOBT.086
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-787-1400
; TELEFAX: 713-787-1440
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 359 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-923-109-2
Query Match 34.5%; Score 609.5; DB 4; Length 359;
Best Local Similarity 40.3%; Pred. No. 2.7e-56;
Matches 139; Conservative 58; Mismatches 109; Indels 39; Gaps 8;

QY 16 AKONSAFPAINCTSETVNAAIKGFADAGSDGIIQFSTGGAEFGSLGVKD-----MV 69
Db 25 AKENNFALPAVNCVGTDSINAVLETAAKVKAPVIVQFSNGGASFIAGKGVKSDVPQGAAI 84
QY 70 TGAVALAEETHVIAAKYPVNVALLTHDCHPKDKLDSVVRPLLATSAQRVSKGNGNPLFQSHM 129
Db 85 LGAISGAHHVHQAHEHYGVPIVILTHDCAK-KLLPWIDGLLDAGEKHFAATGKPLFSSHM 143
```

```
QY 130 WDSAVPIDENLAIAQELKAAAKIILEIEIGVVGGEDGVANE--KLYTSPEDF 187
Db 144 IDLSSESQENIECSKYLERNKSGIMTLEIELGCTGGEEDGVNDSHMDASALYTPEDV 203
QY 188 EKTIBALGAGEHGKYLAAATFGNVHGVKPGNVKLRPDLAQOQVAAAKLGLPADAKPF 247
Db 204 DVAYTEL-SKISPRITAAASFGNVHGVKPGNVVLTPIILRDSQEVVSKHNLPHNS--L 260
QY 248 DFVHGGSGSLKSETEBALRYGVVKNVDTDTQYAFTRPIAGHMTNVDGVULK----- 300
Db 261 NFVHGGSGSTAQEIKDSVSYGVKNVDTDTQWA-----TWEGVLNYYKANE 309
QY 301 -VDGEVG-----VKVYDPRSYLKKAEASMSORVVOACNDLH 336
Db 310 YLQQLGPNKPGEDQPNKKYYDPRVWRAGQTSMIARLEKAFQELN 354

RESULT 7
US-09-934-901-8
; Sequence 8, Application US/09934901
; Patent No. 655353
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odum, J. Martin
; APPLICANT: No. 655353ton, Kelley C.
; APPLICANT: Ye, Rick
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1619 US NA
; CURRENT APPLICATION NUMBER: US/09/934,901
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,906
; PRIOR FILING DATE: September 1, 2000
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 358
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
US-09-934-901-8

Query Match 34.4%; Score 609; DB 4; Length 358;
Best Local Similarity 40.6%; Pred. No. 3e-56;
Matches 145; Conservative 53; Mismatches 121; Indels 38; Gaps 8;

QY 3 IATPEVVAEMLGQAKQNSYAPPAINCTSETVNAIKGFADAGSDGIIQFSTGGAEFGSG 62
Db 12 VVTGEDVQKIFAICKENNFALPAVNVISTDITNAVLEAAAKAKSPVVIQFSNGGAFFVAG 71
QY 63 LGVK-----DMVTGAVALAEFTHVIAAKYPNVVALHTDHCPCDKLDSVVRPLLAIQAORV 117
Db 72 KGLKLEGGCSIHGAISGAHHVRLAELYGVPVVLHTDHAAK-KLLPWVDGMLDEGEKPF 130
QY 118 SKGKNPLFQSHMWDGSAVPIDENLAIAQELKAAAKIILEIEIGVVGGEDGVANE-- 175
Db 131 AATGKPLFSSHMLDLSEESLEENIEICGYLARMAKMGMTLEIELGCTGGEEDGVNDSGM 190
QY 176 INEKLYTSPDEFKTIETALGAGEHGKYLAAATFGNVHGVKPGNVKLRPDLIAOQOQVAA 235
Db 191 DHSALYTQPEDVAYAYEHL-SKISPNFTIAASFGNVHGVSPGNVKLTPEKILDSQKYS 249
QY 236 AKGLPADAKPDEFVHGGSGSLKSEIEEALRYGVVKNVDTDTQYAFTRPIAGHMTN 295
Db 250 EKFGLP--AKSLTFVHGGSGSPPEIKESISYGVVKNVDTDTQWA-----TW 296
QY 296 DGVLK-----VDGEVG-----VKVYDPRSYLKKAEASMSORVVOACNDLH 336
Db 297 EGVNMFYKKNEGYLOQIGNPEGADKPNKKYYDPRVWRAGQEGVVARLQQAFOELN 353

RESULT 8
US-10-321-210-8
; Sequence 8, Application US/10321210
; Patent No. 6767744
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odum, J. Martin
; APPLICANT: No. 6767744ton, Kelley C.
; APPLICANT: Ye, Rick
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1619 US NA
; CURRENT APPLICATION NUMBER: US/10/321,210
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US/09/934,901
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,906
; PRIOR FILING DATE: September 1, 2000
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 358
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
US-10-321-210-8
```

RESULT 8

US-09-934-868-18

; Sequence 18, Application US/09934868

RESULT 11

US-09-248-796A-17059
 ; Sequence 17059, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstock et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDI
 ; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; CURRENT FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 17059

```
;
;      ORGANISM: Candida albicans
US-09-248-796A-17059

Query Match          34.1%; Score 602.5; DB 4; Length 366;
Best Local Similarity 42.9%; Pred. No. 1.5e-55;
Matches 149; Conservative 52; Mismatches 121; Indels 25; Gaps 11
```

Db 261 -TDARKPLVLFPHGSGSTQEEFNATIKNGVVKWNLDTDCQAYLTGIRDIVV-TNKIEYL 318

Qy 300 K--VDGEVGV---KKVYDPRSLLKKAESMSQRVVQACNDLHCAGK 340

Db 319 KAPVGNPEGADPKNKYFDPRVWVREGKTMKSGIAEALDIFHTKQG 365

RESULT 12
US-09-583-110-5316
; Sequence 5316, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:14:10 ; Search time 82.2638 Seconds
(without alignments)
1742.576 Million cell updates/sec

Title: US-10-617-038-4
Perfect score: 1769
Sequence: 1 MPIATPEVVAEMLGQAKONS.....SORVVQACNDLHCAGKSUTH 344

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1769	100.0	344	9	US-09-712-363-165 Sequence 165, App
2	1769	100.0	344	15	US-10-282-122A-62565 Sequence 62565, A
3	1769	100.0	344	15	US-10-282-122A-64372 Sequence 64372, A
4	1769	100.0	344	15	US-10-617-038-4 Sequence 4, Appli
5	1572	88.9	345	15	US-10-282-122A-63684 Sequence 63684, A
6	1347	76.1	344	9	US-09-738-626-6543 Sequence 6543, Ap
7	1347	75.3	344	16	US-10-781-014-60 Sequence 60, Appl
8	1332	75.3	344	15	US-10-282-122A-54046 Sequence 54046, A
9	1234.5	69.8	340	15	US-10-369-493-8299 Sequence 8299, Ap
10	1210.5	68.4	340	14	US-10-156-761-12057 Sequence 12057, A
11	708.5	40.1	359	18	US-10-987-548-5 Sequence 5, Appli

ALIGNMENTS

RESULT 1

US-09-712-363-165
; Sequence 165, Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Sergio H.
; APPLICANT: Marcotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
; FILE REFERENCE: 07419-032001
; CURRENT APPLICATION NUMBER: US/09/712,363
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: PCT/US00/02246
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206,
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 165

```
; ; LENGTH: 344  
; TYPE: PRT  
; ORGANISM: Mycobacterium tuberculosis  
US-09-712-363-165
```

```
Query Match      100.0%; Score 1769; DB 9; Length 344;  
Best Local Similarity 100.0%; Pred. No. 4.3e-156;  
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy   1 MPIATPEVVAEMLGQAOKNSYAFPAINCTSETVNAAIKGFADAGSDGIIOFSTGGAEFG 60  
| | | | |  
Db   1 MPIATPEVVAEMLGQAOKNSYAFPAINCTSETVNAAIKGFADAGSDGIIOFSTGGAEFG 60  
| | | | |  
Qy   61 SGLGVKDMVTGAVALAETHVIATAKYPPNVNLTHTDHCPCDKLDSYVRPLLAISAQRVSKG 120  
| | | | |  
Db   61 SGLGVKDMVTGAVALAETHVIATAKYPPNVNLTHTDHCPCDKLDSYVRPLLAISAQRVSKG 120  
| | | | |  
Qy   121 GNPLFOSHMDGSAPVDENLAIQAELLKAIAAKIILEIEIGVVGEEDGVANEINEKL 180  
| | | | |  
Db   121 GNPLFOSHMDGSAPVDENLAIQAELLKAIAAKIILEIEIGVVGEEDGVANEINEKL 180  
| | | | |  
Qy   181 YTSPEDEFEKTIALGAGEHGKYLAAATFCGNVKLRPDIILAOQQOAAAALGL 240  
| | | | |  
Db   181 YTSPEDEFEKTIALGAGEHGKYLAAATFCGNVKLRPDIILAOQQOAAAALGL 240  
| | | | |  
Qy   241 PADAKPFDFVFHGGSGSLKSEIBEARLYGVVMNDTDTQYAFTPTAGHMFTNYDGVLK 300  
| | | | |  
Db   241 PADAKPFDFVFHGGSGSLKSEIBEARLYGVVMNDTDTQYAFTPTAGHMFTNYDGVLK 300  
| | | | |  
Qy   301 VDGEVGKVYPDRSYLKKEASMSRVSORVVOACNDLHCAGKS LTH 344  
| | | | |  
Db   301 VDGEVGKVYPDRSYLKKEASMSRVSORVVOACNDLHCAGKS LTH 344  
| | | | |
```

```
RESULT 2  
US-10-282-122A-62565  
; Sequence 62565, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09
```

```

; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 64372
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64372

Query Match          100.0%; Score 1769; DB 15; Length 344;
Best Local Similarity 100.0%; Pred. No. 4.3e-156;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPIATPEVVAEMLGQAKQNSYAPPAINCTSSETVNAAIKGFADAGSDGIIQFSTGAEFG 60
DB 1 MPIATPEVVAEMLGQAKQNSYAPPAINCTSSETVNAAIKGFADAGSDGIIQFSTGAEFG 60

QY 61 SGLGVKDMVTGAVALAEFTHVIAAKYPVNVVALHTDHCPCDKLDSYVRPDLAISAQVRVSKG 120
DB 61 SGLGVKDMVTGAVALAEFTHVIAAKYPVNVVALHTDHCPCDKLDSYVRPDLAISAQVRVSKG 120

QY 121 GNPLFQSHMWDGSAVPIDENLAIAOELLKAAAAAKIILIEIGVVGGEEDGVANEINEKL 180
DB 121 GNPLFQSHMWDGSAVPIDENLAIAOELLKAAAAAKIILIEIGVVGGEEDGVANEINEKL 180

QY 181 YTSPEDEFKTIIEALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDLTLAGOQQVAAAKLGL 240
DB 181 YTSPEDEFKTIIEALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDLTLAGOQQVAAAKLGL 240

QY 241 PADAKPFDFVFHGGSGSLKSEIEEALRYGVVKNVVDTDQYAFTRPIAGHMFNTYDGVCLK 300
DB 241 PADAKPFDFVFHGGSGSLKSEIEEALRYGVVKNVVDTDQYAFTRPIAGHMFNTYDGVCLK 300

QY 301 VDGEVGVKKVYDPRSYLKKAEASMSORVVQACNDLHCAGKSLTH 344
DB 301 VDGEVGVKKVYDPRSYLKKAEASMSORVVQACNDLHCAGKSLTH 344

RESULT 4
US-10-617-038-4
; Sequence 4, Application US/10617038
; Publication No. US20040057963A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Peter
; APPLICANT: Rosenkrands, Ida
; APPLICANT: Scrymgeour, Anette
; TITLE OF INVENTION: Therapeutic TB Vaccine
; FILE REFERENCE: SSI5AUSA
; CURRENT APPLICATION NUMBER: US/10/617,038
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: DK PA 2002 01098
; PRIOR FILING DATE: 2002-07-13
; PRIOR APPLICATION NUMBER: US 60/401,725
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 4
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-617-038-4

Query Match          100.0%; Score 1769; DB 15; Length 344;
Best Local Similarity 100.0%; Pred. No. 4.3e-156;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPIATPEVVAEMLGQAKQNSYAPPAINCTSSETVNAAIKGFADAGSDGIIQFSTGAEFG 60
DB 1 MPIATPEVVAEMLGQAKQNSYAPPAINCTSSETVNAAIKGFADAGSDGIIQFSTGAEFG 60

QY 61 SGLGVKDMVTGAVALAEFTHVIAAKYPVNVVALHTDHCPCDKLDSYVRPDLAISAQVRVSKG 120
DB 61 SGLGVKDMVTGAVALAEFTHVIAAKYPVNVVALHTDHCPCDKLDSYVRPDLAISAQVRVSKG 120

QY 121 GNPLFQSHMWDGSAVPIDENLAIAOELLKAAAAAKIILIEIGVVGGEEDGVANEINEKL 180
DB 121 GNPLFQSHMWDGSAVPIDENLAIAOELLKAAAAAKIILIEIGVVGGEEDGVANEINEKL 180

QY 181 YTSPEDEFKTIIEALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDLTLAGOQQVAAAKLGL 240
DB 181 YTSPEDEFKTIIEALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDLTLAGOQQVAAAKLGL 240

QY 241 PADAKPFDFVFHGGSGSLKSEIEEALRYGVVKNVVDTDQYAFTRPIAGHMFNTYDGVCLK 300
DB 241 PADAKPFDFVFHGGSGSLKSEIEEALRYGVVKNVVDTDQYAFTRPIAGHMFNTYDGVCLK 300

QY 301 VDGEVGVKKVYDPRSYLKKAEASMSORVVQACNDLHCAGKSLTH 344
DB 301 VDGEVGVKKVYDPRSYLKKAEASMSORVVQACNDLHCAGKSLTH 344

RESULT 5
US-10-282-122A-63684
; Sequence 63684, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 63684
; LENGTH: 345
; TYPE: PRT
```

ORGANISM: Mycobacterium leprae
US-10-282-122A-63684

Query Match 88.9%; Score 1572; DB 15; Length 345;
Best Local Similarity 87.7%; Pred. No. 1e-137;
Matches 300; Conservative 25; Mismatches 17; Indels 0; Gaps 0;

QY 1 MPIATPEVYAEMLGQAKNSYAFPAINTCTSSSTVNAATKGFADAGSDGIIQFSTGGAERG 60
DB 1 MPIATPEIYAEMLRWAKENSYAFPAINTCTSSSTVNAATKGFADAGSDGIIQFSTGGAERF 60

QY 61 SGLGVKDMVTGAVALAETHVIAAKYPVNVALHTDHCPCPKLDSYVRPDLAISAOVRSGK 120
DB 61 SGLGVKDMVTGAVALAKFTHTIAAKYPINVALHTDHCPCPKLDSYVRPDLAISARRVATG 120

QY 121 GNPLFQSHMWDGSAVPIDENLAIQAELLKAAAAAKIILEIEIGVGVGGBEDGVANEINEL 180
DB 121 KDFLFGSHMWDGSAIPIDENLAIQAELLKAAAAAKIILEIEIGVGVGGBEDGVAGINEKL 180

QY 181 YTSPEDEFKTTIEALGAGEHGKYLAAATGCVNHGVYKPGNVKLRPDLAAGQOQVAAKLG 240
DB 181 YTPKDFVKTIDALGAGEHGKYLAAATGCVNHGVYKPGNVKLRPDLAAGQVAAAKLSQ 240

QY 241 PADAKPFDFVPHGSGSLKSEIEEALRYGVVKNVDTDTQYAFTRPIAGHMFNTYDGVLK 300
DB 241 SEGSRPFDFVPHGSGSEKSEIEEALRYGVVKNVDTDTQYAFTRPVSGHMFNTYDGVLK 300

QY 301 VDGEVGVKKVYDPRSYLKKAESMSQRVVOACNDLHCAGKSL 342
DB 301 VDGVDGNKKVYDPRSYLKKAESMTERVLEACNDLRCAGKSV 342

RESULT 6

US-09-738-626-6543
Sequence 6543, Application US/09738626
Publication No. US20020197605A1

GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: NIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 6543
LENGTH: 344
TYPE: PRT
ORGANISM: Corynebacterium glutamicum

US-09-738-626-6543

Query Match 76.1%; Score 1347; DB 9; Length 344;
Best Local Similarity 75.4%; Pred. No. 9.6e-117;
Matches 257; Conservative 30; Mismatches 54; Indels 0; Gaps 0;

QY 1 MPIATPEVYAEMLGQAKNSYAFPAINTCTSSSTVNAATKGFADAGSDGIIQFSTGGAERG 60
DB 1 MPIATPEVYNEMLDRAKEGGFAPPAINTCTSSSTVNAALKGFAEASDGIQFSTGGAERG 60

QY 61 SGLGVKDMVTGAVALAETHVIAAKYPVNVALHTDHCPCPKLDSYVRPDLAISAOVRSGK 120
DB 61 SGLAVKNKVGAVALAFAFAHAAKSYGINVALHTDHCQKVELDEYVRPDLAISQERVDRG 120

QY 121 GNPLFQSHMWDGSAVPIDENLAIQAELLKAAAAAKIILEIEIGVGVGGBEDGVANEINEL 180
DB 121 ELPLFQSHMWDGSAVPIDENLAIQAELLKAKAANILILEVEIGVGVGGBEDGVAKAGANL 180

QY 181 YTSPEDEFKTTIEALGAGEHGKYLAAATGCVNHGVYKPGNVKLRPDLAAGQOQVAAKLG 240
DB 181 YTSPEDEFKTTDAITGEGKRYLLAAATGCVNHGVYKPGNVKLRPDLAAGQOQVARKLGL 240

QY 241 PADAKPFDFVPHGSGSLKSEIEEALRYGVVKNVDTDTQYAFTRPIAGHMFNTYDGVLK 300
DB 241 ADDALPFDFVPHGSGSEKEIEEALTYGVIKMNVDTTQYAFTRPIVSHMFENTYGVLK 300

QY 301 VDGEVGVKKVYDPRSYLKKAESMSQRVVOACNDLHCAGKS 341
DB 301 IDGEVGNKKAVDPRSYMKKAESMSERIEESCQDLKSVGKT 341

RESULT 7
US-10-781-014-60
Sequence 60, Application US/10781014
Publication No. US20040180408A1

GENERAL INFORMATION:
APPLICANT: Pompejus, Markus
APPLICANT: Kroger, Burkhard
APPLICANT: Schroder, Hartwig
APPLICANT: Zeider, Oskar
APPLICANT: Haberhauer, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
TITLE OF INVENTION: INVOLVED IN CARBON METABOLISM AND ENERGY
TITLE OF INVENTION: PRODUCTION
FILE REFERENCE: BGI-126CPCN
CURRENT APPLICATION NUMBER: US/10/781,014
CURRENT FILING DATE: 2004-02-17
PRIOR APPLICATION NUMBER: US 09/602,740
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/141,031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 60/143,208
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 60/151,572
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19931412.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931413.6
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931419.5
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931420.9
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931424.1
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931428.4
PRIOR FILING DATE: 1999-07-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 784
SEQ ID NO 60
LENGTH: 344
TYPE: PRT
ORGANISM: Corynebacterium glutamicum

US-10-781-014-60

Query Match 76.1%; Score 1347; DB 16; Length 344;
Best Local Similarity 75.4%; Pred. No. 9.6e-117;
Matches 257; Conservative 30; Mismatches 54; Indels 0; Gaps 0;

QY 1 MPIATPEVYAEMLGQAKNSYAFPAINTCTSSSTVNAATKGFADAGSDGIIQFSTGGAERG 60
DB 1 MPIATPEVYNEMLDRAKEGGFAPPAINTCTSSSTVNAALKGFAEASDGIQFSTGGAERG 60

```
QY 61 SGLGVKDMVTGAVALAEFTHTVIAAKYPVNVVALHTDHCPCDKLDSYVRPILAIQAQRVSKG 120
Db 61 SGLAVKXNKVKGAVALLAAFAHEAAKSYGINVALHTDHCQKEVLDYVVRPILAIQAQRVDRG 120
QY 121 GNPLFQSHMWDGSAVPIDENLAIQAELLKAAAAAKIILEIEIGVGGEDGVANEINEKL 180
Db 121 ELPLFQSHMWDGSAVPIDENLAIQAELLKAKAKAANIILEIEIGVGGEDGVAEKAGANL 180
QY 181 YTSPEDEFKTIKALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDILOAQOQVAAAKLGL 240
Db 181 YTSPEDEFKTIKALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDILOAQOQVAAAKLGL 240
QY 241 PADAKPFDVFFHGGSGSLKSEIEEALRYGVVKNVDTDTQYAFTRPIAGHMTFTNYDGVLK 300
Db 241 ADDALPFDVFFHGGSGSEKEIEEALTYGVIKMNVDVDTQYAFTRPIVSHMFTNYDGVLK 300
QY 301 VDGEVGVKKVYDPRSYLKKAASMSORVVOACNDLHCAGKS 341
Db 301 IDGEVGNKKAYDPRSYMKKAEQSMSERIIESCODLKSQVGT 341
```

RESULT 8

```
US-10-282-122A-54046
; Sequence 54046, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
```

```
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
```

```
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54046
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Corynebacterium diptheriae
```

```
US-10-282-122A-54046
```

```
Query Match 75.3%; Score 1332; DB 15; Length 344;
Best Local Similarity 73.5%; Pred. No. 2.4e-115;
```

```
Matches 252; Conservative 34; Mismatches 57; Indels 0; Gaps 0;
QY 1 MPIATPEVVAEMLGOAKQNSYAPPAINTSSTVNAAIKGFADAGSDGIIQFSTGGAERG 60
Db 1 MPIATPEVVAEMLGOAKQNSYAPPAINTSSTVNAAIKGFADAGSDGIIQFSTGGAERG 60
QY 61 SGLGVKDMVTGAVALAEFTHTVIAAKYPVNVVALHTDHCPCDKLDSYVRPILAIQAQRVSKG 120
Db 61 SGLSVKNKVAGACALAAFAHEAAKHYGINVALHTDHCQKEVLDYVVRPILAIQAQRVDRG 120
QY 121 GNPLFQSHMWDGSAVPIDENLAIQAELLKAAAAAKIILEIEIGVGGEDGVANEINEKL 180
Db 121 EQPLFQSHMWDGSAVPIDENLUIAELLKAKAKAHIILEAEIGVGGEDGVAEKAGANL 180
QY 181 YTSPEDEFKTIKALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDILOAQOQVAAAKLGL 240
Db 181 YTSPEDEFKTIKALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDILOAQOQVAAAKLGL 240
QY 241 PADAKPFDVFFHGGSGSLKSEIEEALRYGVVKNVDTDTQYAFTRPIAGHMTFTNYDGVLK 300
Db 241 DDSALPFDVFFHGGSGSEKEIEEALGYGVIKMNVDVDTQYAFTRPVVTHMFTNYDGVLK 300
QY 301 VDGEVGVKKVYDPRSYLKKAASMSORVVOACNDLHCAGKS 343
Db 301 IDGEVGNKKVYDPRSYMKKAEQSMADRIIEACODLHSHVGTLLS 343
```

RESULT 9

```
US-10-369-493-8299
; Sequence 8299, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-1(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 8299
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Thermobifida fusca
```

```
US-10-369-493-8299
```

```
Query Match 69.8%; Score 1234.5; DB 15; Length 340;
Best Local Similarity 69.3%; Pred. No. 2.9e-106;
Matches 237; Conservative 38; Mismatches 64; Indels 3; Gaps 1;
```

```
QY 1 MPIATPEVVAEMLGOAKQNSYAPPAINTSSTVNAAIKGFADAGSDGIIQFSTGGAERG 60
Db 1 MPIATPEVVAEMLGRAKSEGPAFPAINTVSSOTLHAALRGFAEASDGIQVISTGGAERFL 60
QY 61 SGLGVKDMVTGAVALAEFTHTVIAAKYPVNVVALHTDHCPCDKLDSYVRPILAIQAQRVSKG 120
Db 61 SGTITKDMVTGAVALAEFTHTVIAAKYPVNVVALHTDHCPCKEKIDTFVRPILAIQAQRVSKG 120
QY 121 GNPLFQSHMWDGSAVPIDENLAIQAELLKAAAAAKIILEIEIGVGGEDGVANEINEKL 180
Db 121 QEPFLQSHMWDGSAVLEENLQIAELLKESKEARTILEIEIGVGGEDGIVGEINEKL 180
QY 181 YTSPEDEFKTIKALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDILOAQOQVAAAKLGL 240
Db 181 YTTPEGRLTAELVGLGEGYVIAALTFCNVHGVYKPGHVKLRPSVLKEIQDVVGAKYG- 239
QY 241 PADAKPFDVFFHGGSGSLKSEIEEALRYGVVKNVDTDTQYAFTRPIAGHMTFTNYDGVLK 300
```

```
Db      240 --KEXPFDLVHFHGGSSLEEIEHAIISYGVVKNMIDTDTQVAFTRPIVDHIMKNYDGLK 297
QY      301 VDGEVGVKKVYDPRSYLKKAEASMSQVRVQACNDLHCAGKSL 342
Db      298 VDGVDGNKAYDPRSYGKAAEAGMARVVEAAQQLKSAGKKL 339

RESULT 10
US-10-156-761-12057
; Sequence 12057, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12057
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12057

Query Match      68.4%; Score 1210.5; DB 14; Length 340;
Best Local Similarity 69.0%; Pred. No. 5e-104;
Matches 236; Conservative 39; Mismatches 64; Indels 3; Gaps 1;

QY      1 MPIATPEVVAEMLGQAKNSYAFFAINCTSSSTVNAAIKGFADAGSDGIIQFSTGAEFG 60
Db      1 MPIATPEVNVNMLDRAKAGKAFYAPAINVTSTQTLHAALRGFAEBSDGIQVISTGAEFL 60

QY      61 SGLGVKDMVTGAVALAEFTHVIAAKYPVNVALHTDHCPCDKLDSYVRPLLAISAQRVSKG 120
Db      61 GQYKKDMVTGAVALAEFAHIAEKYPTVVALHTDHCPCDKLGDYVRPLLAISBERRVAG 120

QY      121 GNPLFQSHMWDGSAVPIDENLAIAQELLKAAAKAIIIEIEIGVVGGEEDGVANIEKL 180
Db      121 RNPLFQSHMWDGSAETLADNLAIAQELLEKARAAKIIIEVEITPTGGBEDGVTHEINDSL 180

QY      181 YTSPEDEFTIEALGAGEHGKYLAAATGNVHVYKPGNVKLRPDLAQQOQVAAKGL 240
Db      181 YTVVEDAIRTAELGUGEGRYLLAASFGNVHVYKPGNVVLRPELLKELNEGVAAKFG- 239

QY      241 PADAKPFDVFFHGGSSLSKSEIEEALRYGVVKNVNDTDTQVAFTRPIAGHMFNTYDGLK 300
Db      240 --KAQPFDFVHGGSGSTEARTALENGVVKMNDTDTQVAFTRPVADHMFNRYDGLK 297

QY      301 VDGEVGVKKVYDPRSYLKKAEASMSQVRVQACNDLHCAGKSL 342
Db      298 VDGEVGDKKYDPTWTKGLAEASMAKRVTEACANLRSTGTGI 339

RESULT 11
US-10-987-548-5
; Sequence 5, Application US/10987548
; Publication No. US20050130280A1
; GENERAL INFORMATION:
; APPLICANT: E. I. DuPont de Nemours & Co., Inc.
; APPLICANT: Pollak, Dana M. Walters
; APPLICANT: Zhu, Quinn
; TITLE OF INVENTION: FRUCTOSE-BISPHOSPHATE ALDOLASE REGULATORY SEQUENCES FOR GENE
; TITLE OF INVENTION: FRUCTOSE-BISPHOSPHATE ALDOLASE REGULATORY SEQUENCES FOR GENE
```

```
; FILE REFERENCE: CL2411
; CURRENT APPLICATION NUMBER: US/10/987,548
; CURRENT FILING DATE: 2004-11-12
; PRIOR APPLICATION NUMBER: US 60/519971
; PRIOR FILING DATE: 2003-11-14
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Pasteurella multocida (GenBank Accession NP_246800)
US-10-987-548-5

Query Match      40.1%; Score 708.5; DB 18; Length 359;
Best Local Similarity 44.0%; Pred. No. 3.5e-57;
Matches 158; Conservative 58; Mismatches 104; Indels 39; Gaps 9;

QY      3 IATPEVVAEMLGQAKNSYAFFAINCTSSSTVNAAIKGFADAGSDGIIQFSTGAEFGSG 62
Db      12 VVTGDDVQKVAYAKANNFAIPAVNCVGSDSVNAVLETAARVKAPVLIQFSNGGAQFVAG 71

QY      63 LGVMDM-----VTGAVALAEFTHVIAAKYPVNVALHTDHCPCDKLDSYVRPLLAISAQR 116
Db      72 KGLXPASGARTDVLGAJAGAKVHALABEYGPVILHTDHAAK-KLLPWIDGLLEAGEEH 130

QY      117 VSKGNPLFQSHMWDGSAVPIDENLAIAQELLKAAAKAIIIEIEIGVVGGEEDGVAN-E 175
Db      131 PAETGKPLFSSHMDLSEEPMEENNAICREVLARMDKMGMTLEIEIGITGGEEDGVNDSD 190

QY      176 INE-KLYTSPEDFEKTIIEALGAGEHGKYLAAATGNVHVYKPGNVKLRPDLAQQOQVA 234
Db      191 VEESKLYTQPEDVLYVVDQLNP-VSPRTVAAAFGNVHVYKPGNVKLPKPSILGASQEFV 249

QY      235 AAKLGLPADAKPFDVFFHGGSSLSKSEIEEALRYGVVKNVNDTDTQVAFTRPIAGHMFNTN 294
Db      250 SKERGLP--AKSIDFVPHGGSGSREIREAISYGAIKMNIDTDTOWA-----S 296

QY      295 YDGLVK-----VDGEVG-----VKKVYDPRSYLKKAEASMSQVRVQACNDLJHC 337
Db      297 WEGILKFYNANQDYLQOLGNPEGPDAPNKKYDPRVWLKRWESMSKRLSQSFEDLNC 355

RESULT 12
US-09-815-242-11040
; Sequence 11040, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
```

```

; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 1410
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11040
; LENGTH: 359
; TYPE: PRF
; ORGANISM: Haemophilus influenzae
; US-09-815-242-11040

Query Match      38.4%; Score 679.5; DB 9; Length 359;
Best Local Similarity 42.5%; Pred. No. 1.8e-54;
Matches 154; Conservative 61; Mismatches 102; Indels 45; Gaps 10;

QY 3 IATPEVYAEMLCQAKNSYAFPAINCTSETVNAAIKGFADAGSDGIIQFSTGGAEFGSG 62
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
12 VVTGEDVQKVFAYAKEHNFPAIVNCGSDSVNAVLETAARVKAPVIIQFSNGGAIFYAG 71
QY 63 LGVKDM-----VTGAVALAEFTHTAAKYPVNVVALHTDCHPKDKLDSVVRPLLAISAQR 116
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
72 KGIKPTSGTRPDVLGAIAAGAKQVHTLAKEYGVPVILTDHAAK-KLLPWIDGLLDAGEKH 130
QY 117 VSKGNPLFQSHMWDGSAVPIDENLAIAOELLKAAAAAKIILEIEIGVVGGEDGVAN-E 175
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
131 FAETGRPLFSSHMIDLSEESMEENMAICREYLARMKMGMTLEIEIGITGGEDGVNDSD 190
QY 176 INE-KLYTSPDEKTEIALGAGEH---GKYLAAATFGNVHGVYKPGNVKLRPDILAQOQ 231
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
191 VDERSLYTQPSDLYVYDQL-----HPVSPNFTVAAAFGNVHGVYKPGNVKLRKPSILGESQ 246
QY 232 QVAAAKLGLPADAKPFDVPHGSGSLKSEIEBEALRYGVVKNVDTDTQYAFTRPIAGHM 291
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
247 EFVSKERNLP--AKPINFVPHGSGSSREIEAIGYGAIKWNIDTDTQWA-----295
QY 292 FTNYDGVLK-----VDGEVG-----VKVYDPRSYLKKABASMSQVVOACNDL 335
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
296 --SWNGILNFYKANEAYLQQLGNPEGPDAPNKKYYDPRVWLKMEESMSKLEQSFEDL 353
QY 336 HC 337
Db :
354 NC 355

RESULT 13
US-10-282-122A-58200
; Sequence 58200, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09

```

```

; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58200
; LENGTH: 359
; TYPE: PRF
; ORGANISM: Haemophilus influenzae
; US-10-282-122A-58200

Query Match      38.4%; Score 679.5; DB 15; Length 359;
Best Local Similarity 42.5%; Pred. No. 1.8e-54;
Matches 154; Conservative 61; Mismatches 102; Indels 45; Gaps 10;

QY 3 IATPEVYAEMLCQAKNSYAFPAINCTSETVNAAIKGFADAGSDGIIQFSTGGAEFGSG 62
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
12 VVTGEDVQKVFAYAKEHNFPAIVNCGSDSVNAVLETAARVKAPVIIQFSNGGAIFYAG 71
QY 63 LGVKDM-----VTGAVALAEFTHTAAKYPVNVVALHTDCHPKDKLDSVVRPLLAISAQR 116
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
72 KGIKPTSGTRPDVLGAIAAGAKQVHTLAKEYGVPVILTDHAAK-KLLPWIDGLLDAGEKH 130
QY 117 VSKGNPLFQSHMWDGSAVPIDENLAIAOELLKAAAAAKIILEIEIGVVGGEDGVAN-E 175
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
131 FAETGRPLFSSHMIDLSEESMEENMAICREYLARMKMGMTLEIEIGITGGEDGVNDSD 190
QY 176 INE-KLYTSPDEKTEIALGAGEH---GKYLAAATFGNVHGVYKPGNVKLRPDILAQOQ 231
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
191 VDERSLYTQPSDLYVYDQL-----HPVSPNFTVAAAFGNVHGVYKPGNVKLRKPSILGESQ 246
QY 232 QVAAAKLGLPADAKPFDVPHGSGSLKSEIEBEALRYGVVKNVDTDTQYAFTRPIAGHM 291
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
247 EFVSKERNLP--AKPINFVPHGSGSSREIEAIGYGAIKWNIDTDTQWA-----295
QY 292 FTNYDGVLK-----VDGEVG-----VKVYDPRSYLKKABASMSQVVOACNDL 335
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
296 --SWNGILNFYKANEAYLQQLGNPEGPDAPNKKYYDPRVWLKMEESMSKLEQSFEDL 353
QY 336 HC 337
Db :
354 NC 355

RESULT 14
US-10-987-548-4
; Sequence 4, Application US/10987548
; Publication No. US20050130280A1
; GENERAL INFORMATION:
; APPLICANT: E. I. DuPont de Nemours & Co., Inc.
; APPLICANT: Pollak, Dana M. Walters
; APPLICANT: Zhu, Quinn
; TITLE OF INVENTION: FRUCTOSE-BISPHOSPHATE ALDOLASE REGULATORY SEQUENCES FOR GENE
; FILE REFERENCE: CL2411
; CURRENT APPLICATION NUMBER: US/10/987,548
; CURRENT FILING DATE: 2004-11-12
; PRIOR APPLICATION NUMBER: US 60/519971
; PRIOR FILING DATE: 2003-11-14
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 359
; TYPE: PRF
; ORGANISM: Haemophilus influenzae (GenBank Accession NP_438682)

```


Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	1769	100.0	344	2	D70576	probable fructose	
2	1572	88.9	345	2	S69444	probable fructose	
3	1342	75.9	344	2	S09283	fructose-bisphosp	
4	1212	68.5	343	2	T36539	probable fructose	
5	679.5	38.4	359	2	C64074	fructose-bisphosp	
6	661	37.4	358	2	G82317	fructose-bisphosp	
7	636	36.0	359	2	D70155	fructose-bisphosp	
8	635.5	35.9	354	2	S52413	fructose-bisphosp	
9	626.5	35.4	359	2	A10112	fructose-bisphosp	
10	609.5	34.5	359	1	ADECA2A	fructose-bisphosp	
11	609.5	34.5	359	2	D91103	fructose-bisphosp	
12	609.5	34.5	384	2	G85948	fructose-bisphosp	
13	603.5	34.1	359	2	AC0875	fructose-1,6-bisph	
14	598.5	33.8	358	2	E84982	fructose-bisphosp	
15	571.5	32.3	358	2	T39798	fructose-bisphosp	
16	569.5	32.2	358	2	T43289	fructose-bisphosp	
17	566.5	32.0	366	2	D47260	fructose-bisphosp	
18	531	30.0	359	1	ADBY2	fructose-bisphosp	
19	329.5	18.6	288	2	E54202	fructose-bisphosp	
20	327.5	18.5	288	2	S73455	fructose-bisphosp	
21	327	18.5	293	2	S97938	fructose-bisphosp	
22	327	18.5	293	2	D95070	fructose-bisphosp	
23	320.5	18.1	298	2	E86863	fructose-bisphosp	
24	307.5	17.4	286	2	F90525	fructose-bisphosp	
25	302	17.1	286	2	B30006	fructose-bisphosp	
26	297.5	16.8	284	2	AD1394	fructose-1,6-bisph	
27	297.5	16.8	284	2	AG1769	fructose-1,6-bisph	
28	295.5	16.7	347	2	S85949	probable fructose	
29	292	16.5	284	2	AH0103	probable aldolase	

A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-344 <COL>
A: Cross-references: UNIPROT: O06313; GB: Z95324; GB: AL123456; NID: g3261760; PIDN: CAB08571.
A: Experimental source: strain H37Rv
C: Genetics:
A: Gene: fba
C: Superfamily: fructose-bisphosphate aldolase II

A:Gene: lba
C:Superfamily: fructose-bisphosphate aldolase II

	Query Match	100.0%;	Score 1769;	DB 2;	Length 344;
	Best Local Similarity	100.0%;	Pred. No. 7e-122;	Mismatches 0;	Indels 0; Gaps 0;
	Matches 344;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1 MPIATPEVYAEMLGAKONSYAFPAINCSTSETVNAAIKGFADAGSDGIIFQTGGAEFG	60			
Dd	1 MPIATPEVYAEMLGAKONSYAFPAINCSTSETVNAAIKGFADAGSDGIIFQTGGAEFG	60			
Qy	61 SGLGVKDMVTGAVALAETHVIAAKYPNVNVALHTDHCPKDKLDSYVRPLLAISAQRVSKG	120			
Dd	61 SGLGVKDMVTGAVALAETHVIAAKYPNVNVALHTDHCPKDKLDSYVRPLLAISAQRVSKG	120			
Qy	121 GNPLFQSHMWDCGSVPIDENLAIAOELLKAAAAAKIILEIEIGVGWGEEDGVANEINEKL	180			
Dd	121 GNPLFQSHMWDCGSVPIDENLAIAOELLKAAAAAKIILEIEIGVGWGEEDGVANEINEKL	180			
Qy	181 YTSPEDEFTKTIALGAGBHGKYLLAATFGNVHGVYKPGNVKLRPDILAOQQVAAAALGL	240			
Dd	181 YTSPEDEFTKTIALGAGBHGKYLLAATFGNVHGVYKPGNVKLRPDILAOQQVAAAALGL	240			
Qy	241 PADAKPPDFVHHGSGSLKSEIEEALRYGVVKMNVDTDTQYAFTRPFIAGHMFTNYDGVLK	300			
Dd	.241 PADAKPPDFVHHGSGSLKSEIEEALRYGVVKMNVDTDTQYAFTRPFIAGHMFTNYDGVLK	300			
Qy	301 VDGEVGVKKVDPRSYLKAEBASMSQRVVQAACNDLHCACKSLTH	344			
Dd	301 VDGEVGVKKVDPRSYLKAEBASMSQRVVQAACNDLHCACKSLTH	344			

RESULT 5

C64074
fructose-bisphosphate aldolase (EC 4.1.2.13) II - Haemophilus influenzae (strain Rd KW20)
C/Species: Haemophilus influenzae
C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C/Accession: C64074
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
J.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
.D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A.;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;/Reference number: A64000; MUID:95350630; PMID:7542800
A/Accesion: C64074
A;/Status: nucleic acid sequence not shown; translation not shown
A;/Molecule type: DNA
A;/Residues: 1-359 <TIG>
A;/Cross-references: UNIPROT:P44429; GB:U32734; GB:I42023; NID:g1573498; PIDN:AAC22182.1;
C/Superfamily: fructose-bisphosphate aldolase II
C;/Keywords: aldehyde-lyase; phosphate-carbon lyase; gluconeogenesis; glycolysis; pentose ph

Query Match 38.4%; Score 679.5; DB 2; Length 359;
Best Local Similarity 42.5%; Pred. No. 3.3e+42;
Matches 154; Conservative 61; Mismatches 102; Indels 45; Gaps 10;

Qy 3 IATPEVVAEMLGQAQNYSYAPPAINCTSTSETNAAIKGFADAGSDGTIOFTSGAEFGSG 62
Db :|::||:::||:::||::||::||::||::||::||::||::||::||::||::||
Qy 12 VVTGEDVOKVFAYAKEHNFAIPAVNCVGSDSVNAVILETAARVKAPVIIQFSNGGAIFYAG 71
Db :|::||:::||:::||::||::||::||::||::||::||::||::||::||::||
Qy 63 LGVKDM-----VTGAVALAEFTHTIAAKYPNVNLTHDHCPKDLDVSVPRLAISAQK 116
Db :|::||:::||:::||::||::||::||::||::||::||::||::||::||::||
Qy 72 KGIKPTSGTRPDVLGALAGAKQVHTLAKEGVGPVILHDDAAK-KLLPMWIDLGLAGEKH 130
Db :|::||:::||:::||::||::||::||::||::||::||::||::||::||::||
Qy 117 VSKGNPLFOSHMDGSAPIDENLATAOELLKAAAANKIIIEIETGVGGEDGYAN-E 175
Db :|::||:::||:::||::||::||::||::||::||::||::||::||::||::||
Qy 131 FAETGRFLPSHMIDUSEESMEENMAICREYLARMDKMGMTLEIEGITGEEDGDVNSD 190
Db :|::||:::||:::||::||::||::||::||::||::||::||::||::||::||
Qy 176 INE-KLYTSPEDFEKTTEALGAGEH---GKYLLAATFGNVHVYKPGNKVKLRPDIILAOGQ 231
Db :|::||:::||:::||::||::||::||::||::||::||::||::||::||::||
Qy 191 VDSESLYTQSFDLVIVYDQL----HPVSPNFTVAALAFGNVHGPKGNVKLFPSILGESQ 246
Db :|::||:::||:::||::||::||::||::||::||::||::||::||::||::||
Qy 232 QVAAKLGLPADAKPDFVFPHGSGSLKSRIEALRYGVVMKNVDITDYAFTRPIAGHM 291
Db :|::||:::||:::||::||::||::||::||::||::||::||::||::||::||
Qy 247 EFVSKERNLP-AKPINFVPHGSGSSREEIREAIGYGAIKMVIDDTQWA----- 295
Db :|::||:::||:::||::||::||::||::||::||::||::||::||::||::||
Qy 292 FTNYDGVLK-----VDGEVG-----VKKVDPDRSYLKAASMSQRVVQCNDL 335
Db :|::||:::||:::||::||::||::||::||::||::||::||::||::||::||
Qy 296 --SWNGILNFYKANEAAYLOGQLNPGEPPAPNNKKYYDPRVWLKRMBEESMKRLEQSFEDL 353
Db :|::||:::||:::||::||::||::||::||::||::||::||::||::||::||
Qy 336 HC 337
Db :|::||:::||:::||::||::||::||::||::||::||::||::||::||::||
Qy 354 NC 355
Db :|::||:::||:::||::||::||::||::||::||::||::||::||::||::||

RESULT 6

G82317
fructose-bisphosphate aldolase, class II VC0478 [imported] - Vibrio cholerae (strain N16)

C/Species: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C/Accession: G82317
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vanathavan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
Nature 406, 477-483, 2000
A;/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;/Reference number: A82035; MUID:20406833; PMID:10952301
A/Accesion: G82317
A;/Status: preliminary
A;/Molecule type: DNA
A;/Residues: 1-358 <HEI>
A;/Cross-references: UNIPROT:Q9KUN7; GB:AB0004134; GB:AB0003852; NID:g9654900; PIDN:AAF9336

A:Accession: A38058
A:Molecule type: protein
A:Residues: 165-167;213-216;277-285;337-345 <ALE2>
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, D.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:978503
A:Accession: D65077
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-359 <BLAT>
A:Cross-references: GB:AE000376; GB:U00096; NID:g2367176; PIDN:AACT5962.1; PID:g1789293;
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: fba; fda
A:Map position: 63 min
C:Complex: homodimer [validated, MUID:89193446]
C:Function:
A:Description: EC 4.1.2.13 [validated, MUID:89193446]
A:Note: class II enzymes utilize a divalent metal ion to act as the electron sink
C:Superfamily: fructose-bisphosphate aldolase II
C:Keywords: aldehyde-lyase; carbon-carbon lyase; gluconeogenesis; glycolysis, homodimer;
F.2-359/Product: fructose-bisphosphate aldolase #status experimental <MAT>

Query Match 34.5%; Score 609.5; DB 1; Length 359;
Best Local Similarity 40.3%; Pred. No. 4.3e-37;
Matches 139; Conservative 58; Mismatches 109; Indels 39; Gaps 8;

QY 16 AKONSAYFAINCTSETVNAAIKGFADAGSDGIIOFSTGGAEFGSGLVKQD-----MV 69
DB 25 AKENNFALPAVNCVGTDSINAVLEAAKVPVIVQFSGNGASFIAGKGVKSDVPQGAAI 84
70 TGAVALAEETHVIAAKYPVNVALHTDHCPCPKLDLSVVRPLLAISAOVRKSGGNPLFQSHM 129
85 LGAISGAHHVQMAEHYGVPIVILHTDCAK-KLLPWIDGLLDAGEKHFAATGKPLFSSHM 143
130 WDGSAVPIDENIAIAQELKAAAKAIIIEIGVVGGEEDGVANEINE--KLYTSPEDF 187
144 IDLSEESLQENIEICKYLERMSKIGMTLEIELGCTGGEEDGVNDSHMDASALYTQPEDV 203
188 EKTIEALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDLIAOQOQVAAAKGLGPADAKPF 247
204 DYAYTEL-SKISPRFTIAASFGNVHGVYKPGNVLTPTILRDSQEVYVSKKHNLPHNS--L 260
248 DFVFHGGSGSLKSEIEALRYGVVKNVDTDTQYAFTRPIAGHMFNTYDGVLK----- 300
261 NFVFHGGSGSTAQEIKDSVSYGVVKNIDTDTQWA-----TWEGVLNYYKANE 309

Query Match 34.5%; Score 609.5; DB 1; Length 359;
Best Local Similarity 40.3%; Pred. No. 4.3e-37;
Matches 139; Conservative 58; Mismatches 109; Indels 39; Gaps 8;

QY 16 AKONSAYFAINCTSETVNAAIKGFADAGSDGIIOFSTGGAEFGSGLVKQD-----MV 69
DB 25 AKENNFALPAVNCVGTDSINAVLEAAKVPVIVQFSGNGASFIAGKGVKSDVPQGAAI 84
70 TGAVALAEETHVIAAKYPVNVALHTDHCPCPKLDLSVVRPLLAISAOVRKSGGNPLFQSHM 129
85 LGAISGAHHVQMAEHYGVPIVILHTDCAK-KLLPWIDGLLDAGEKHFAATGKPLFSSHM 143
130 WDGSAVPIDENIAIAQELKAAAKAIIIEIGVVGGEEDGVANEINE--KLYTSPEDF 187
144 IDLSEESLQENIEICKYLERMSKIGMTLEIELGCTGGEEDGVNDSHMDASALYTQPEDV 203
188 EKTIEALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDLIAOQOQVAAAKGLGPADAKPF 247
204 DYAYTEL-SKISPRFTIAASFGNVHGVYKPGNVLTPTILRDSQEVYVSKKHNLPHNS--L 260
248 DFVFHGGSGSLKSEIEALRYGVVKNVDTDTQYAFTRPIAGHMFNTYDGVLK----- 300
261 NFVFHGGSGSTAQEIKDSVSYGVVKNIDTDTQWA-----TWEGVLNYYKANE 309

Query Match 34.5%; Score 609.5; DB 2; Length 384;
Best Local Similarity 40.3%; Pred. No. 4.7e-37;
Matches 139; Conservative 58; Mismatches 109; Indels 39; Gaps 8;

QY 16 AKONSAYFAINCTSETVNAAIKGFADAGSDGIIOFSTGGAEFGSGLVKQD-----MV 69
DB 50 AKENNFALPAVNCVGTDSINAVLEAAKVPVIVQFSGNGASFIAGKGVKSDVPQGAAI 109
70 TGAVALAEETHVIAAKYPVNVALHTDHCPCPKLDLSVVRPLLAISAOVRKSGGNPLFQSHM 129
110 LGAISGAHHVQMAEHYGVPIVILHTDCAK-KLLPWIDGLLDAGEKHFAATGKPLFSSHM 168
130 WDGSAVPIDENIAIAQELKAAAKAIIIEIGVVGGEEDGVANEINE--KLYTSPEDF 187
169 IDLSEESLQENIEICKYLERMSKIGMTLEIELGCTGGEEDGVNDSHMDASALYTQPEDV 228
188 EKTIEALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDLIAOQOQVAAAKGLGPADAKPF 247
229 DYAYTEL-SKISPRFTIAASFGNVHGVYKPGNVLTPTILRDSQEVYVSKKHNLPHNS--L 285
248 DFVFHGGSGSLKSEIEALRYGVVKNVDTDTQYAFTRPIAGHMFNTYDGVLK----- 300

C:Superfamily: fructose-bisphosphate aldolase II
Query Match 34.5%; Score 609.5; DB 2; Length 359;
Best Local Similarity 40.3%; Pred. No. 4.3e-37;
Matches 139; Conservative 58; Mismatches 109; Indels 39; Gaps 8;

QY 16 AKONSAYFAINCTSETVNAAIKGFADAGSDGIIOFSTGGAEFGSGLVKQD-----MV 69
DB 25 AKENNFALPAVNCVGTDSINAVLEAAKVPVIVQFSGNGASFIAGKGVKSDVPQGAAI 84
70 TGAVALAEETHVIAAKYPVNVALHTDHCPCPKLDLSVVRPLLAISAOVRKSGGNPLFQSHM 129
85 LGAISGAHHVQMAEHYGVPIVILHTDCAK-KLLPWIDGLLDAGEKHFAATGKPLFSSHM 143
130 WDGSAVPIDENIAIAQELKAAAKAIIIEIGVVGGEEDGVANEINE--KLYTSPEDF 187
144 IDLSEESLQENIEICKYLERMSKIGMTLEIELGCTGGEEDGVNDSHMDASALYTQPEDV 203
188 EKTIEALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDLIAOQOQVAAAKGLGPADAKPF 247
204 DYAYTEL-SKISPRFTIAASFGNVHGVYKPGNVLTPTILRDSQEVYVSKKHNLPHNS--L 260
248 DFVFHGGSGSLKSEIEALRYGVVKNVDTDTQYAFTRPIAGHMFNTYDGVLK----- 300
261 NFVFHGGSGSTAQEIKDSVSYGVVKNIDTDTQWA-----TWEGVLNYYKANE 309

Query Match 34.5%; Score 609.5; DB 2; Length 384;
Best Local Similarity 40.3%; Pred. No. 4.7e-37;
Matches 139; Conservative 58; Mismatches 109; Indels 39; Gaps 8;

QY 16 AKONSAYFAINCTSETVNAAIKGFADAGSDGIIOFSTGGAEFGSGLVKQD-----MV 69
DB 50 AKENNFALPAVNCVGTDSINAVLEAAKVPVIVQFSGNGASFIAGKGVKSDVPQGAAI 109
70 TGAVALAEETHVIAAKYPVNVALHTDHCPCPKLDLSVVRPLLAISAOVRKSGGNPLFQSHM 129
110 LGAISGAHHVQMAEHYGVPIVILHTDCAK-KLLPWIDGLLDAGEKHFAATGKPLFSSHM 168
130 WDGSAVPIDENIAIAQELKAAAKAIIIEIGVVGGEEDGVANEINE--KLYTSPEDF 187
169 IDLSEESLQENIEICKYLERMSKIGMTLEIELGCTGGEEDGVNDSHMDASALYTQPEDV 228
188 EKTIEALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDLIAOQOQVAAAKGLGPADAKPF 247
229 DYAYTEL-SKISPRFTIAASFGNVHGVYKPGNVLTPTILRDSQEVYVSKKHNLPHNS--L 285
248 DFVFHGGSGSLKSEIEALRYGVVKNVDTDTQYAFTRPIAGHMFNTYDGVLK----- 300

```

Db 286 NFVFHGGSGTAQAEIKDSVGVVKKMNDTDTQWA-----TWEGVLNYYKANE 334
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 301 -VDGEVG-----VKKVYDPRSLYKKAASMSQVVQACNDLH 336
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 335 YLQGLGNPKGEDQPNKKYDPRVWLRAGQTSMIARLEKAFQELN 379
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 13
AC0875
fructose 1,6-bisphosphate aldolase [imported] - Salmonella enterica sero
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AC0875
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AC0875
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-359 <PAR>
A:Cross-references: GB:ALU513382; PIDN:CAD02899.1; PID:gl6504152; GSPDB:GN00176
C:Genetics:
C:Gene: STY3226
C:Superfamily: fructose-bisphosphate aldolase II

Query Match 34.1%; Score 603.5; DB 2; Length 359;
Best Local Similarity 40.3%; Pred. No. 1.2e-36;
Matches 139; Conservative 56; Mismatches 111; Indels 39; Gaps 8;

Qy 16 AKONSYAPPAINCTSSETVNAAIKGFADAGSDGIIOFSTGGAEGFSGGLGVKD-----MV 69
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 25 AKENFALPANCVCCTDSINAVLETAARKVPIVQFNSNGASFIAGKGVKTDVPOGA 84
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 70 TGAVALAEFTHVIAAKYPVNVVALHTDHCPCPKLDSYVRPLLAISAQRVSKGNPLFQSHM 129
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 85 LCAISGAHHVQMAEYGVGVLTHTDCAK-KLLPWIDGLLDAGEKHPAATGKPLFSSHM 143
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 130 WDGSAPVIDENLAIAQELLKAAAKAIIIEIGVVGGEEDGVANE--KLYTSPEDF 187
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 144 IDLSSESHENIEICSKYLARMSKIGMTLEIELGCTGGEEDGVDSHMDASALYTPEDV 203
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 188 EKTIEALGAGEHGKYLAAATFGNVHGVKPGNVKLRPDILAQQGVAAAKLGLPADAKPF 247
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 204 DYAYTEL-SKISPRFTIAASFGNVHGVKPGNVVLTPTILRDSQDYVSKHNLPHNS--L 260
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 248 DFVPHGGSGSLKSEIEEALRYGVVKKVNVVDTDTQYAFTRPIAGHMTNVDYGVLK----- 300
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 261 NFVFHGGSGTAQAEIKDSVGVVKKMNDTDTQWA-----TWEGVLNYYKANE 309
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 301 -VDGEVG-----VKKVYDPRSLYKKAASMSQVVQACNDLH 336
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 310 YLQGLGNPKGEDQPNKKYDPRVWLRAGQTSMIARLEKAFKELN 354
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 14
E84982
fructose-bisphosphate aldolase (EC 4.1.2.13) [imported] - Buchnera sp. (strain APS)
C:Species: Buchnera sp.
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: E84982
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A
A:Reference number: AB4930; MUID:20445173; PMID:10993077
A:Accession: E84982
A:Status: preliminary
A:Molecule type: DNA
```

```

A:Residues: 1-358 <STO>
A:Cross-references: GB:AP000398; GSPDB:GN00144
A:Experimental source: strain APS
C:Genetics:
A:Gene: fba, BU451
C:Superfamily: fructose-bisphosphate aldolase II
C:Keywords: aldehyde-lyase; carbon-carbon lyase
```

```

Query Match 33.8%; Score 598.5; DB 2; Length 358;
Best Local Similarity 39.2%; Pred. No. 2.7e-36;
Matches 133; Conservative 65; Mismatches 114; Indels 27; Gaps 8;

Qy 16 AKONSYAPPAINCTSSETVNAAIKGFADAGSDGIIOFSTGGAEGFSGGLGVK-----DW 69
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 24 AKKQFALPANCVCCTDSINAVLETAARKVPSPIIOFSHGGASFIAGYKKLSENOEAI 83
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 70 TGAVALAEFTHVIAAKYPVNVVALHTDHCPCPKLDSYVRPLLAISAQRVSKGNPLFQSHM 129
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 84 QGAVSQAQVHLMAKHVEIPVLHTDHCPCKELL-SWIDGLLEVGQNYFYNNRPLFTSHM 142
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 130 WDGSAPVIDENLAIAQELLKAAAKAIIIEIGVVGGEEDGVAN-EINPK-LYTSPEDF 187
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 143 IDLSKESLEENISTCKYFKRIKNNMMELEIELGCTGGEEDGIONTKIDKKLLYTPQDV 202
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 188 EKTIEALGAGEHGKYLAAATFGNVHGVKPGNVKLRPDILAQQGVAAAKLGLPADAKPF 247
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 203 NYAYEELNTISK-NFSTAASFGNVHGVQPNIDRLPILKNSQEFVSSKHNL--EKNPL 259
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 248 DFVPHGGSGSLKSEIEEALRYGVVKKVNVVDTDTQYAFTRPIAG-----HMFTNYDG 297
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 260 NLVPHGGSGSLKEIKESIQGVVKKMNDITDQIAAWKGVLDIFYKQNKFEFLQHLGHTTN 319
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 298 VLKVDGEVGVKVPDPSYLYKKAASMSQVVQACNDLH 336
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 320 KHKPN-----KKYDPRTRWIRKQESISIRLEQSFKELN 353
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 15

```

T39798
fructose-bisphosphate aldolase (EC 4.1.2.13) - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T39798
R:McDougall, R.C.; Rajandram, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.
submitted to the EMBL Data Library, July 1999
A:Reference number: Z21880
A:Accession: T39798
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-358 <MCD>
A:Cross-references: UNIPROT:P36580; EMBL:AL109731; PIDN:CAB52034.1; GSPDB:GN00067; SPDB:
A:Experimental source: strain 972h-; cosmid c19C2
C:Genetics:
A:Gene: SPDB:SPBC19C2.07
A:Map position: 2
C:Superfamily: fructose-bisphosphate aldolase II
C:Keywords: aldehyde-lyase; carbon-carbon lyase
```

```

Query Match 32.3%; Score 571.5; DB 2; Length 358;
Best Local Similarity 39.0%; Pred. No. 2.6e-34;
Matches 135; Conservative 58; Mismatches 138; Indels 15; Gaps 6;
```

```

Qy 3 IATEVVAEMLGQAKQNSYAPAINCTSSSETVNAAIKGFADAGSDGIIOFSTGGAEGFSG 62
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 11 VITGDNVLKLFYARBHGFAIPAINVTSSSTAIAALEAAREARSPILLOTNGGAHFFAG 70
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 63 L-----GVKDMVTGAVALAEFTHVIAAKYPVNVVALHTDHCPCPKLDSYVRPLLAISAQRV 117
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 71 KESNEGOKASIAAGIAAAHYIRSIAPFFGVVWVHSDHCAK-KLLPMDCGMFDEAYF 129
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 118 SKGNPLFQSHMDGSAVPIDENLAIAQELLKAAAKAIIIEIGVVGGEEDGVANE-- 175
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 130 KIHGEPFLSSHMLDLSEEPKKENIAQVKEYCKRAVPMKIWIEMETGITGGBEDGVDSNVH 189
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:24:25 ; Search time 77.0538 Seconds
(without alignments)
2286.133 Million cell updates/sec

Title: US-10-617-038-4
Perfect score: 1769
Sequence: 1 MPIATPEVVAEMLGQAKQNS.....SORVVQANDLHCAGKSLTH 344

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1769	100.0	344	1	ALF_MYCBO	P67476 mycobacteri
2	1769	100.0	344	1	ALF_MYCTU	P67475 mycobacteri
3	1572	88.9	345	1	ALF_MYCLE	O69600 mycobacteri
4	1372	77.6	344	2	O8FMA6	O8fma6 corynebacte
5	1342	75.9	343	1	ALF_CORGL	P19537 corynebacte
6	1332	75.3	344	2	O6NF15	O6nf15 corynebacte
7	1247.5	70.5	340	2	O70AQ6	O70aq6 nondumarea
8	1212	68.5	343	1	ALF_STRCO	O9x8r6 streptomyc
9	1210.5	68.4	340	1	ALF_STRGB	O9zem7 streptomyc
10	1210.5	68.4	340	2	O8XEU0	O82eu0 streptomyc
11	1195.5	67.6	341	2	O6A687	O6a687 propionibac
12	1101	62.2	355	2	O8G6T8	O8g6t8 bifidobacte
13	1089	61.6	341	2	O6ADV3	O6adv3 leifsonia x
14	708.5	40.1	359	2	O9CJX8	O9cjx8 pasteurella
15	697.5	39.4	358	2	O7VMV1	O7vmv1 haemophilus
16	697.5	39.4	359	2	O6SW09	O6sw09 manheimia
17	679.5	38.4	359	1	ALF_HAEIN	P44429 haemophilus
18	668	37.8	358	2	O87LL2	O87ll2 vibrio para
19	661	37.4	358	2	O9KUN7	O9kun7 vibrio chol
20	651	36.8	439	2	O84X64	O84x64 phaeodactyl
21	650	36.7	358	2	O7N7Z4	O7n7z4 photorhabdu
22	646	36.5	358	1	ALF_EDWIC	O52402 edwardsiell
23	638	36.1	358	2	O8GF88	O8gf88 photorhabdu
24	636	36.0	359	1	ALF_BORBU	O51401 borrelia bu
25	635.5	35.9	354	1	ALF_CAWJE	P53818 campylobact
26	635	35.9	358	2	O6D088	O6d088 erwinia car
27	632	35.7	358	2	O6LNM2	O6lmn2 photobacter
28	626.5	35.4	359	2	O8ZHH4	O8zhh4 yersinia pe
29	626	35.4	359	2	O661H5	O661h5 borrelia ga
30	624.5	35.3	342	2	O666Q2	O666q2 yersinia ps
31	621	35.1	402	2	O8L882	O8l882 odontella s

Query Match 100.0%; Score 1769; DB 1; Length 344;

RESULT 1

ALF_MYCBO
ID ALF_MYCBO STANDARD; PRT; 344 AA.
AC P67476: O06313;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 25-OCT-2004 (Rel. 45, Last annotation update)
DE Fructose-bisphosphate aldolase (EC 4.1.2.13).
GN Name=fba; OrderedLocusNames=Mb0370c;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eigelmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7897-7882(2003).
CC -!- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycero
CC phosphate + D-glyceraldehyde 3-phosphate.
CC -!- COFACTOR: Zinc (By similarity).
CC -!- PATHWAY: Glycolysis; sixth step.
CC -!- SIMILARITY: Belongs to the class II fructose-bisphosphate aldolase
family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: BX248335; CAD93233.1; --
HSSP: P11604; 1ZEN
InterPro: IPR006411; Fruct_bisp_bact.
DR InterPro; IPR000771; K_bp_aldolase.
DR Pfam; PF01116; F_bp_aldolase; 1.
DR ProDom; PD002376; K_bp_aldolase; 1.
DR TIGRFAMs; TIGR00167; cba; 1.
DR TIGRFAMs; TIGR01520; FruBisaldo II A; 1.
DR PROSITE; PS00502; ALDOLASE_CLASS_II_1; 1.
DR PROSITE; PS00806; ALDOLASE_CLASS_II_2; 1.
KW Complete proteome; Glycolysis; Lyase; Zinc.
FT METAL 93 93 Zinc (By similarity).
FT METAL 96 96 Zinc (By similarity).
SQ SEQUENCE 344 AA; 36544 MW; 04FA3E1123F72FB1 CRC64;

```
Best Local Similarity 100.0%; Pred. No. 2.8e-116; Mismatches 0; Indels 0; Gaps 0;
Matches 344; Conservative 0;

QY 1 MPIATPEVYAEMLGQAKNSYAFPAINTSETVNAAIKGFADAGSDGIIQFSTGGAEFG 60
DB 1 MPIATPEVYAEMLGQAKNSYAFPAINTSETVNAAIKGFADAGSDGIIQFSTGGAEFG 60
QY 61 SGLGVKDMVTGAVALAEETHVIAAKYPVNVALHTDHCPCDKLDSVVRPLLAISQORVSKG 120
DB 61 SGLGVKDMVTGAVALAEETHVIAAKYPVNVALHTDHCPCDKLDSVVRPLLAISQORVSKG 120
QY 121 GNPLFQSHMDGSAVPIDENLAIAQELLKAAAKIIIEIIGVVGGEEDGVANEINEKL 180
DB 121 GNPLFQSHMDGSAVPIDENLAIAQELLKAAAKIIIEIIGVVGGEEDGVANEINEKL 180
QY 181 YTSPEDEFEKTIEALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDIILAQOQVAAAKLGL 240
DB 181 YTSPEDEFEKTIEALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDIILAQOQVAAAKLGL 240
QY 241 PADAKPDPFVPHGGSGSLKSEIEEALRYGVVKNMVDTDQYAFTRPIAGHMFTNYDGVK 300
DB 241 PADAKPDPFVPHGGSGSLKSEIEEALRYGVVKNMVDTDQYAFTRPIAGHMFTNYDGVK 300
QY 301 VDGEVGVKKVYDPRSYLKKAEASMSQORVQACNDLHCAGKSLTH 344
DB 301 VDGEVGVKKVYDPRSYLKKAEASMSQORVQACNDLHCAGKSLTH 344

RESULT 2
ALF_MYCTU STANDARD; PRT; 344 AA.
ID ALF_MYCTU
AC P67475; O06313;
DT 30-MAY-2000 (Rel. 39, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DE Fructose-bisphosphate aldolase (EC 4.1.2.13).
DE Name=fba; OrderedLocusNames=Rv0363c, MT0379; ORFNames=MTCV13E10.25c;
OS Mycobacterium tuberculosis
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=11773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekaiia F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Uterback T.R., Weidman J.P., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains";
RL J. Bacteriol. 184:5479-5490(2002).
CC -!- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycerone
CC phosphate + D-glyceraldehyde 3-phosphate.
CC -!- COPACTOR: Zinc (By similarity).
```

```
-!- PATHWAY: Glycolysis; sixth step.
-!- SIMILARITY: Belongs to the class II fructose-bisphosphate aldolase
family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; BX842573; CAB08571.1; -
EMBL; AE000516; AAK44600.1; -
PIR; D70576; D70576.
HSSP; P11604; 1ZEN.
TIGR; MT0379; -
TubercuList; Rv0363c; -
InterPro; IPR006411; Fruct_bisp_bact.
InterPro; IPR000771; K_bp_aldolase.
Pfam; PF01116; E_bp_aldolase; 1.
PRODOM; PD002376; K_bp_aldolase; 1.
DR TIGRFAMs; TIGR00167; cbbA; 1.
DR TIGRFAMs; TIGR01520; FruBisAldo_II_A; 1.
DR PROSITE; PS00602; ALDOLASE_CLASS_II_1; 1.
DR PROSITE; PS00806; ALDOLASE_CLASS_II_2; 1.
KW Complete proteome; Glycolysis; Lyase; zinc.
FT METAL 93 Zinc (By similarity).
FT METAL 96 Zinc (By similarity).
SQ SEQUENCE 344 AA; 36544 MW; 04FA3E1123F72FB1 CRC64;

Query Match 100.0%; Score 1769; DB 1; Length 344;
Best Local Similarity 100.0%; Pred. No. 2.8e-116;
Matches 344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPIATPEVYAEMLGQAKNSYAFPAINTSETVNAAIKGFADAGSDGIIQFSTGGAEFG 60
DB 1 MPIATPEVYAEMLGQAKNSYAFPAINTSETVNAAIKGFADAGSDGIIQFSTGGAEFG 60
QY 61 SGLGVKDMVTGAVALAEETHVIAAKYPVNVALHTDHCPCDKLDSVVRPLLAISQORVSKG 120
DB 61 SGLGVKDMVTGAVALAEETHVIAAKYPVNVALHTDHCPCDKLDSVVRPLLAISQORVSKG 120
QY 121 GNPLFQSHMDGSAVPIDENLAIAQELLKAAAKIIIEIIGVVGGEEDGVANEINEKL 180
DB 121 GNPLFQSHMDGSAVPIDENLAIAQELLKAAAKIIIEIIGVVGGEEDGVANEINEKL 180
QY 181 YTSPEDEFEKTIEALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDIILAQOQVAAAKLGL 240
DB 181 YTSPEDEFEKTIEALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDIILAQOQVAAAKLGL 240
QY 241 PADAKPDPFVPHGGSGSLKSEIEEALRYGVVKNMVDTDQYAFTRPIAGHMFTNYDGVK 300
DB 241 PADAKPDPFVPHGGSGSLKSEIEEALRYGVVKNMVDTDQYAFTRPIAGHMFTNYDGVK 300
QY 301 VDGEVGVKKVYDPRSYLKKAEASMSQORVQACNDLHCAGKSLTH 344
DB 301 VDGEVGVKKVYDPRSYLKKAEASMSQORVQACNDLHCAGKSLTH 344

RESULT 3
ALF_MYCLE
ID ALF_MYCLE STANDARD; PRT; 345 AA.
AC O69600;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Fructose-bisphosphate aldolase (EC 4.1.2.13).
DE Name=fba; OrderedLocusNames=ML0286; ORFNames=MLCBA.29c;
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
```

```

RN  SEQUENCE FROM N.A.
RC  STRAIN=TN;
RX  MEDLINE=21128732; PubMed=11234002; DOI=10.1038/35059006;
RA  Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA  Wheeler P.R., Honore N., Garnier T., Churcher C.M., Harris D.E.,
RA  Mungall K.L., Basham D., Brown D., Chillingworth T., Connor R.,
RA  Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA  Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA  Murphy L.D., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA  Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA  Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA  Barrell B.G.;
RT  "Massive gene decay in the leprosy bacillus.";
RL  Nature 409:1007-1011(2001).
CC  -!- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycerone
CC  phosphate + D-glyceraldehyde 3-phosphate.
CC  -!- COFACTOR: Zinc (By similarity).
CC  -!- PATHWAY: Glycolysis; sixth step.
CC  -!- SIMILARITY: Belongs to the class II fructose-bisphosphate aldolase
CC  family.
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
DR  EMBL; AL023514; CAA18950.1; -.
DR  EMBL; AL583918; CAC29794.1; -.
DR  PIR; F86944; F86944.
DR  HSP; P11604; I1857.
DR  Leptoma; ML0286; -.
DR  InterPro; IPR006411; Fruct_bisp_bact.
DR  InterPro; IPR000771; K_bp_aldolase.
DR  Pfam; PF01116; F_bp_aldolase; 1.
DR  ProDom; PD002376; K_bp_aldolase; 1.
DR  TIGRfam; TIGR00167; cbbA; 1.
DR  TIGRfam; TIGR01520; FruBisAldo II A; 1.
DR  PROSITE; PS00602; ALDOLASE_CLASS_II_1; 1.
DR  PROSITE; PS00806; ALDOLASE_CLASS_II_2; 1.
KW  Complete proteome; Glycolysis; lyase; Zinc.
FT  METAL 93 93 Zinc (By similarity).
FT  METAL 96 96 Zinc (By similarity).
SQ  SEQUENCE 345 AA; 36740 MW; 29F88474C41D00C7 CRC64;
-----
Query Match 88.9%; Score 1572; DB 1; Length 345;
Best Local Similarity 87.7%; Pred. No. 2e-102;
Matches 300; Conservative 25; Mismatches 17; Indels 0; Gaps 0;

QY 1 MPIATPEVVAEMLGQAKNSYAFPAINCTSSTVNAAIKGFADAGSDGIIQSTGGAEPG 60
DB 1 MPIATPEIIVAEMLRRAKENSYAFPAINCTSSTVNAAIKGFADAGSDGIIQSTGGAEPG 60
QY 61 SGLGVKDMVTGAVALAEFTHTVIAAKYPVNVALTHDCHPCPKLDSDYVVRPLLAI SARVATG 120
DB 61 SGLGVKDMVTGAVALAEFTHTVIAAKYPVNVALTHDCHPCPKLDSDYVVRPLLAI SARVATG 120
QY 121 GNPLFQSHMWDGSAVPIDENLAIAOELLKAAAKIILEIEIGVVGGEEDGVANEINEKL 180
DB 121 KDFLFGSHMWDGSAIPIDENLAIAOELLKAAAKIILEIEIGVVGGEEDGVANEINEKL 180
QY 181 YTPSPDFEKTIALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDLTLAGQOQVAAAKLGL 240
DB 181 YTPSPDFEKTIALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDLTLAGQOQVAAAKLGL 240
QY 241 PADAKPFPDFVFGGSGSLKSEIEEALRYGVYKMNVDTDTOYAFTRPIAGHMFNTYDGLVK 300
DB 241 SEGSKFPDFVFGGSGSLKSEIEEALRYGVYKMNVDTDTOYAFTRPIAGHMFNTYDGLVK 300
QY 301 VDEGVGVKKVYDPRSYLKKAEASMSORVVOACNDLHCAGKSV 342

```

```

DB 301 VDEGVGVKKVYDPRSYLKKAEASMSORVVOACNDLHCAGKSV 342
-----
RESULT 4
O8FMA6 PRELIMINARY; PRT; 344 AA.
ID O8FMA6;
AC O8FMA6;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Fructose-bisphosphate aldolase (EC 4.1.2.13).
GN Name=fda; OrderedLocusNames=CE2601;
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]_TaxID=152794;
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314;
RX MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens.";
RL Genome Res. 13:1572-1579(2003).
CC -!- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycerone
CC phosphate + D-glyceraldehyde 3-phosphate.
CC -!- COFACTOR: Zinc (By similarity).
CC -!- PATHWAY: Glycolysis; sixth step.
CC -!- SIMILARITY: Belongs to the class II fructose-bisphosphate aldolase
CC family.
DR EMBL; AF005223; BAC19411.1; -.
DR HSP; P11604; IZEN.
DR GO; GO:0004332; F:fructose-bisphosphate aldolase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006411; Fruct_bisp_bact.
DR InterPro; IPR000771; K_bp_aldolase.
DR Pfam; PF01116; F_bp_aldolase; 1.
DR ProDom; PD002376; K_bp_aldolase; 1.
DR TIGRfam; TIGR00167; cbbA; 1.
DR TIGRfam; TIGR01520; FruBisAldo II A; 1.
DR PROSITE; PS00602; ALDOLASE_CLASS_II_1; 1.
DR PROSITE; PS00806; ALDOLASE_CLASS_II_2; 1.
KW Complete proteome; Glycolysis; Lyase.
SQ SEQUENCE 344 AA; 37189 MW; 6B2F0F077322C35C CRC64;
-----
Query Match 77.6%; Score 1372; DB 2; Length 344;
Best Local Similarity 77.1%; Pred. No. 2.3e-88;
Matches 263; Conservative 29; Mismatches 49; Indels 0; Gaps 0;

QY 1 MPIATPEVVAEMLGQAKNSYAFPAINCTSSTVNAAIKGFADAGSDGIIQSTGGAEPG 60
DB 1 MPIATPEVVAEMLGQAKNSYAFPAINCTSSTVNAAIKGFADAGSDGIIQSTGGAEPG 60
QY 61 SGLGVKDMVTGAVALAEFTHTVIAAKYPVNVALTHDCHPCPKLDSDYVVRPLLAI SARVATG 120
DB 61 SGLGVKDMVTGAVALAEFTHTVIAAKYPVNVALTHDCHPCPKLDSDYVVRPLLAI SARVATG 120
QY 121 GNPLFQSHMWDGSAVPIDENLAIAOELLKAAAKIILEIEIGVVGGEEDGVANEINEKL 180
DB 121 GNPLFQSHMWDGSAVPIDENLAIAOELLKAAAKIILEIEIGVVGGEEDGVANEINEKL 180
QY 181 YTPSPDFEKTIALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDLTLAGQOQVAAAKLGL 240
DB 181 YTPSPDFEKTIALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDLTLAGQOQVAAAKLGL 240
QY 241 PADAKPFPDFVFGGSGSLKSEIEEALRYGVYKMNVDTDTOYAFTRPIAGHMFNTYDGLVK 300
DB 241 PADAKPFPDFVFGGSGSLKSEIEEALRYGVYKMNVDTDTOYAFTRPIAGHMFNTYDGLVK 300

```

```

Db 241 ADDALPFDVFGGSGSEKEIEALRYGVIKMNVDTDTQVAFTRPIASHMFENVGVK 300
QY 301 VDGEVGVKKVDPDPSRYLKAESMSQRVVQACNDLHCAGKS 341
Db 301 IDGEVGNKAYDPRSYLKAESQMSERVIESQDLHSGVGT 341

RESULT 5
ALP_CORGL STANDARD; PRT; 343 AA.
AC P19537;
DT 01-NOV-1990 (Rel. 16, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE Fructose-bisphosphate aldolase (EC 4.1.2.13).
GN Name=fba; Synonyms=fda; OrderedLocusNames=Cgl2770, cg3068;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
SEQUENCE FROM N.A., AND SEQUENCE OF 1-10.
RC STRAIN=ATCC 13059 / AS019;
RX MEDLINE=90136092; PubMed=2615658;
RA von der Osten C.H., Barbas C.F. III, Wong C.-H., Sinskey A.J.;
RT "Molecular cloning, nucleotide sequence and fine-structural analysis
of the Corynebacterium glutamicum fda gene: structural comparison of
C. glutamicum fructose-1,6-bisphosphate aldolase to class I and class
II aldolases.";
RL Mol. Microbiol. 3:1625-1637(1989).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RN [3]
SUBMITTED (MAY-2002) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;
RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
RA Burkovski A., Dusch N., Eggelein L., Eikmanns B.J., Gaigalat L.,
RA Goessmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
RA McHardy A.C., Meyer F., Moockel B., Pfeifferle W., Puhler A.,
RA Rey D.A., Rueckert C., Rupp O., Sam H., Wendisch V.F., Wiegand I.,
RA Tauch A.;
RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
and its impact on the production of L-aspartate-derived amino acids
and vitamins.";
RL J. Biotechnol. 104:5-25(2003).
CC -!- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycerone
phosphate + D-glyceraldehyde 3-phosphate.
CC -!- COFACTOR: Zinc.
CC -!- PATHWAY: Glycolysis; sixth step.
CC -!- SIMILARITY: Belongs to the class II fructose-bisphosphate aldolase
family.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; X17313; CAA35190.1; -.
DR EMBL; AP005282; BAC00164.1; -.
DR EMBL; BX927156; CAF20791.1; -.
DR PIR; S09283; S09283.
DR HSRF; F11604; IZEN.
DR InterPro; IPR006411; Fruct_bisp_bact.
DR InterPro; IPR000771; K_bp_aldolase.
DR Pfam; PF01116; F_bp_aldolase; 1.

```

```

DR ProDom: PD002376; K_bp_aldolase; 1.
DR TIGRFAMs; TIGR00167; c5bA; 1.
DR TIGRFAMs; TIGR01520; fruBisAldo_II_A; 1.
DR TIGRFAMs; TIGR01859; fruc_bis_ald; 1.
DR PROSITE; PS00602; ALDOLASE_CLASS_II_1; 1.
DR PROSITE; PS00806; ALDOLASE_CLASS_II_2; 1.
KW Complete proteome; Direct protein sequencing; Glycolysis; Lyase; Zinc.
FT INIT_MET 0
FT METAL 92 Zinc (By similarity).
FT METAL 95 Zinc (By similarity).
FT CONFLICT 279 Q -> H (in Ref. 1).
SQ SEQUENCE 343 AA; 37083 MW; F8D5A12F416CF35 CRC64;

Query Match 75.9%; Score 1342; DB 1; Length 343;
Best Local Similarity 75.3%; Pred. No. 2.9e-86;
Matches 256; Conservative 30; Mismatches 54; Indels 0; Gaps 0;

QY 2 PIATPEVVAEMLGQAKNSVAFPAINCTSSETVNAAIKGFADAGSDGIIQFSTGAEFGS 61
Db 1 PIATPEVYNEMLDRAKEGGFAFPAINCTSETINAAIKGFADAGSDGIIQFSTGAEFGS 60

QY 62 GLGVKDMVTGAVALAEFTHVIAAKYPVYNVALHTDCHPKDLDSYVRPLLAISAQRVSKG 121
Db 61 GLAVKNVKGAVALLAFAHEAAKSYGINVALHTDCHQKVELDDEVVRPLLAISQERVDGR 120

QY 122 NPLFQSHMWGDSAVPIDENLAIQELLKAAAAAKIILEIGTVGVGGEDGVANENHKLY 181
Db 121 LPLFQSHMWGDSAVPIDENLEIAQELLAKAKAANIILEVEIGTVGVGGEDGVAKAGANLY 180

QY 182 TSPEDFEKTEALCAGEHGKYLAAATFGNVHGVYKPGNVKLRPDLAQGQOAAAKLGLP 241
Db 181 TSPEDFEKTDAGTGEKGYLLAATFGNVHGVYKPGNVKLRPEVLLEGQOVARKKLGLA 240

QY 242 ADAKPFDFVHGGSGSLKSEIEALRYGVVKMNVDTDTQVAFTRPIACHMFTNVDGLVK 301
Db 241 DDALPFDVFGGSGSEKEIEALRYGVIKMNVDTDTQVAFTRPIVSHMPENYNGVLKI 300

QY 302 DGEVGVKKVYDPRSYLKAESMSQRVVQACNDLHCAGKS 341
Db 301 DGEVGNKAYDPRSYLKAESQMSERVIESQDLHSGVGT 340

RESULT 6
Q6NF15 Q6NF15 PRELIMINARY; PRT; 344 AA.
AC Q6NF15;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Fructose-bisphosphate aldolase (EC 4.1.2.13)
GN Name=fba; Synonyms=fda; OrderedLocusNames=DIP2094;
OS Corynebacterium diphtheriae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1717;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Biotype Gravis / NCTC 13129;
RX MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/gkg874;
RA Cerdeno-Tarraga A.-M., Efstratiou A., Dover L.G., Holden M.T.G.,
RA Pallen M.J., Bentley S.D., Besra G.S., Churcher C.M., James K.D.,
RA De Zoraya A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
RA Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,
RA Rabinowitz E., Rutherford K.M., Thomson N.R., Unwin L.,
RA Whitehead S., Barrall B.G., Parkhill J.;
RT "The complete genome sequence and analysis of Corynebacterium
diphtheriae NCTC13129.";
RL Nucleic Acids Res. 31:6516-6523(2003).
CC -!- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycerone
phosphate + D-glyceraldehyde 3-phosphate.
CC -!- COFACTOR: Zinc (By similarity).
CC -!- PATHWAY: Glycolysis; sixth step.
CC -!- SIMILARITY: Belongs to the class II fructose-bisphosphate aldolase
family.

```

```

CC EMBL; BX248360; CAB50622.1; -.
DR HSSP; P11604; 1B57.
DR GO; GO:0004332; F:fructose-bisphosphate aldolase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006411; Fruct_bisp_bact.
DR InterPro; IPR000771; K_bp_aldolase.
DR Pfam; PF01116; F_bp_aldolase; 1.
DR ProDom; PD002376; K_bp_aldolase; 1.
DR TIGRFAMs; TIGR00167; cbbA; 1.
DR TIGRFAMs; TIGR01520; FruBisaldo II A; 1.
DR PROSITE; PS00602; ALDOLASE_CLASS_II_1; 1.
DR PROSITE; PS00806; ALDOLASE_CLASS_II_2; 1.
KW Complete proteome; Glycolysis; Lyase.
SQ SEQUENCE 344 AA; 37256 MW; BDFBF2B9BC538260 CRC64;

Query Match 75.3%; Score 1332; DB 2; Length 344;
Best Local Similarity 73.5%; Pred. No. 1.5e-85;
Matches 252; Conservative 34; Mismatches 57; Indels 0; Gaps 0;

QY 1 MPIATPEVVAEMLGQAKNSYAPPAINCTSSTVNAAIKGFADAGSDGIIQFSTGGAERG 60
DB 1 MPIATPEVVAEMLGQAKNSYAPPAINCTSSTVNAAIKGFADAGSDGIIQFSTGGAERG 60
QY 61 SGLGVKDMVTGAVALAEFTHVIAAKYPVNVVALHTDHCPCDKLDSYVRPLAISAQVRSG 120
DB 61 SGLSVKNVAGACALAAFAHEAKHYGINVALTHDHCQKLEVDYVRPLAISQVRDRG 120
QY 121 GNPLFQSHWWDGSAVPIDENLAIAQELLKAAAKAIIIEIGVGVGEGEDGVANEINEKL 180
DB 121 EQPLFQSHWWDGSAIPIDENLVIAQELLKAKAKAIIIEAIEIGVGVGEGEDGVAKAGANL 180
QY 181 YTSPEDEKTIETALGAGEHGKYLAAATFGNVHGVKPGNVKLRPDLAQQQVAAAKLGL 240
DB 181 YTSPEDEKTIETALGAGEHGKYLAAATFGNVHGVKPGNVKLRPDLAQQQVAAAKLGL 240
QY 241 PADAKPFDVFHFGSGSLKSEIEALRYGVVKNVNDTDTQYAFTRPIAGHMFTNYDGLVK 300
DB 241 DDSALPFDVFHFGSGSEKIEEALGYGVKNVNDTDTQYAFTRPIAGHMFTNYDGLVK 300
QY 301 VDGEVGVKKVYDPRSYLKAEASMSQVRVQACNDLHCAGKSLT 343
DB 301 IDGEVGNKKTYPDSYKMAEQAMADRIIEACODLHSGVTLS 343

RESULT 7
Q70AQ6 PRELIMINARY; PRT; 340 AA.
AC Q70AQ6;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Putative fructose-1,6-bisphosphate aldolase (EC 4.1.2.13).
GN Name=fda;
OS Nonomurea sp. ATCC 39727.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptosporangineae; Streptosporangiaceae; Nonomurea.
OX NCBI_TaxID=93944;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15101992; DOI=10.1111/j.1365-2958.2004.04028.x;
RA Gunnarsson N., Mortensen U.H., Sosio M., Nielsen J.;
RT Identification of the Entner-Doudoroff pathway in an antibiotic
RL Mol. Microbiol. 52:895-902 (2004).
CC -1- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycero-
CC phosphate + D-glyceraldehyde 3-phosphate.
CC -1- COFACTOR: Zinc (By similarity).
CC -1- PATHWAY: Glycolysis; sixth step.
CC -1- SIMILARITY: Belongs to the class II fructose-bisphosphate aldolase
CC family.

```

```

DR EMBL; AJ605557; CAB53637.2; -.
DR HSSP; P11604; 1B57.
DR GO; GO:0004332; F:fructose-bisphosphate aldolase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006411; Fruct_bisp_bact.
DR InterPro; IPR000771; K_bp_aldolase.
DR Pfam; PF01116; F_bp_aldolase; 1.
DR ProDom; PD002376; K_bp_aldolase; 1.
DR TIGRFAMs; TIGR00167; cbbA; 1.
DR TIGRFAMs; TIGR01520; FruBisaldo II A; 1.
DR PROSITE; PS00602; ALDOLASE_CLASS_II_1; 1.
DR PROSITE; PS00806; ALDOLASE_CLASS_II_2; UNKNOWN_1.
KW Glycolysis; Lyase.
SQ SEQUENCE 340 AA; 36410 MW; 31F9D6AC524F323 CRC64;

Query Match 70.5%; Score 1247.5; DB 2; Length 340;
Best Local Similarity 69.7%; Pred. No. 1.3e-79;
Matches 239; Conservative 41; Mismatches 60; Indels 3; Gaps 1;

QY 1 MPIATPEVVAEMLGQAKNSYAPPAINCTSSTVNAAIKGFADAGSDGIIQFSTGGAERG 60
DB 1 MPIATPEVVAEMLGQAKNSYAPPAINCTSSTVNAAIKGFADAGSDGIIQFSTGGAERG 60
QY 61 SGLGVKDMVTGAVALAEFTHVIAAKYPVNVVALHTDHCPCDKLDSYVRPLAISAQVRSG 120
DB 61 SGATIKDMVTGATALAEYARVVAKYPTVALHTDHCPCDKLDDGFMRLIDISLERVSG 120
QY 121 GNPLFQSHWWDGSAVPIDENLAIAQELLKAAAKAIIIEIGVGVGEGEDGVANEINEKL 180
DB 121 LDPLFQSHWWDGSAVPIDENLEIAKELLKDCARAKIIMEIEIGVGVGEGEDGVANEINEKL 180
QY 181 YTSPEDEKTIETALGAGEHGKYLAAATFGNVHGVKPGNVKLRPDLAQQQVAAAKLGL 240
DB 181 YTTADALATAEAVGVGDKGRYMLAATFGNVHGVKPGNVKLRPDLAQQQVAAAKLGL 240
QY 241 PADAKPFDVFHFGSGSLKSEIEALRYGVVKNVNDTDTQYAFTRPIAGHMFTNYDGLVK 300
DB 240 --KDKPFDLVFHHGSGSLKEIEQAEISYGVVKNVNDTDTQYAFTRPIAGHMFTNYDGLVK 297
QY 301 VDGEVGVKKVYDPRSYLKAEASMSQVRVQACNDLHCAGKSLT 343
DB 298 IDGEVGNKKTYPDSYKMAEQAMARVLEACQSLKSAGTKIS 340

RESULT 8
ALF_STRCO STANDARD; PRT; 343 AA.
AC Q9XBR6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Fructose-bisphosphate aldolase (EC 4.1.2.13).
GN Name=fba; OrderedLocusNames=SCO3649; ORFNames=SCH10.27c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Baccan A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares S., Taylor K.,
RA Warren T., Wierzbicki A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces

```

```

RT coelcolor A3(2)";
RL Nature 417,141-147(2002).
CC -!- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycero-
CC phosphate + D-glycerolaldehyde 3-phosphate.
CC -!- COFACTOR: Zinc (By similarity).
CC -!- PATHWAY: Glycolysis; sixth step.
CC -!- SIMILARITY: Belongs to the class II fructose-bisphosphate aldolase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL939117; CAB42036.1; -.
CC PIR; T36539; T36539.
CC DR HSP; P11604; 1B57.
CC DR InterPro; IPR006411; Fruct_bisp_bact.
CC DR InterPro; IPR000771; K_bp_aldolase.
CC DR Pfam; PF01116; F_bp_aldolase; 1.
CC DR ProDom; PD002376; K_bp_aldolase; 1.
CC DR TIGRFAMs; TIGR00167; cbaA; 1.
CC DR TIGRFAMs; TIGR01520; FruBisAldo II A; 1.
CC DR PROSITE; PS00802; ALDOLASE_CLASS_II_1; 1.
CC DR PROSITE; PS00806; ALDOLASE_CLASS_II_2; FALSE NEG.
CC KW Complete proteome; Glycolysis; Lyase; Zinc.
CC FT METAL 93 93 Zinc (By similarity).
CC FT METAL 96 96 Zinc (By similarity).
CC SQ SEQUENCE 343 AA; 36926 MW; A1E2BE768028E2C CRC64;

Query Match 68.5%; Score 1212; DB 1; Length 343;
Best Local Similarity 67.8%; Pred. No. 4e-77;
Matches 232; Conservative 43; Mismatches 67; Indels 0; Gaps 0;

QY 1 MPIATPEVYAEMLGQAKQNSYAFPAINCTSETVNAAIKGFADAGSDGIIFSTGGAEFG 60
DB 1 MPIATPEVYNEMLDRAKAGKAFYPAINTVTSQTALRALRFAEASDGIQISTGAEFL 60

QY 61 SGLGVKDMVTGAVALAEETHVIAAKYPNVVALHTDHCPCDKLDSVVRPLLAI SAQRVSKG 120
DB 61 GQHNKDMVTGAVALAEFAHIAEKYDVTVALHTDHCPCDKLDSVVRPLIAVSEERVAG 120

QY 121 GNPLFQSHMWDGSAVPIDENLAIQAELLKAAAKIILEIEIGVVGGEDGVANEINEKL 180
DB 121 RNPLFQSHMWDGSAETLADNLISIAQELLARARARAIILEVEITPTGGEDGVSHINDSL 180

QY 181 YTSPEDEKTIETALGAGEHGKYLAAATFGNVHGVKPGNVKLRPDLIAQGOQVAAAKLGL 240
DB 181 YTTVDADVKTETALGAGEGKRYLLAASFGNVHGVKPGNVKLRPDLIAQGOQVAAAKLGL 240

QY 241 PADAKPFDVFHGGSGSLKSEIEEALRYGVVKMNVDTDTQYAFTRPIAGHMTNVDGVLK 300
DB 241 PAGESKPFDFVHGGSGSTAEETATALEGVVKMNVDTDTQYAFTRPVVDHFRNVDGVLK 300

QY 301 VDGEVGVKKVYDPRSYLKKAASMSQRVVQACNDLHCAGKSL 342
DB 301 VDGEVGNKKTYPDRTWGLAEAGMAARVVEACGHLRSAGQKI 342

RESULT 9
ALF_STRGB STANDARD; PRT; 340 AA.
AC Q9ZEM7;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Fructose-bisphosphate aldolase (EC 4.1.2.13) (FBP aldolase).
GN Name=fba; Synonyms=fda;
OS Streptomyces galbus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

```

```

Streptomycineae; Streptomycetaceae; Streptomyces.
NCBI_TaxID=33898;
OX [1]
RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=DSM 40480;
RX MEDLINE=21184406; PubMed=11287146;
RA Wehmeier U.F.;
RT "Molecular cloning, nucleotide sequence and structural analysis of the
RT Streptomyces galbus DSM40480 fda gene: the S. galbus fructose-1,6-
RT bisphosphate aldolase is a member of the class II aldolases.";
RL FEMS Microbiol. Lett. 197:53-58(2001).
CC -!- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycero-
CC phosphate + D-glycerolaldehyde 3-phosphate.
CC -!- COFACTOR: Zinc.
CC -!- PATHWAY: Glycolysis; sixth step.
CC -!- MISCELLANEOUS: Its optimum pH is 7.5.
CC -!- SIMILARITY: Belongs to the class II fructose-bisphosphate aldolase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ131707; CA10483.2; -.
CC DR HSP; P11604; 12EN.
CC DR InterPro; IPR006411; Fruct_bisp_bact.
CC DR InterPro; IPR000771; K_bp_aldolase.
CC DR Pfam; PF01116; F_bp_aldolase; 1.
CC DR ProDom; PD002376; K_bp_aldolase; 1.
CC DR TIGRFAMs; TIGR00167; cbaA; 1.
CC DR TIGRFAMs; TIGR01520; FruBisAldo II A; 1.
CC DR PROSITE; PS00802; ALDOLASE_CLASS_II_1; 1.
CC DR PROSITE; PS00806; ALDOLASE_CLASS_II_2; FALSE NEG.
CC KW Glycolysis; Lyase; Zinc.
CC FT METAL 93 93 Zinc (By similarity).
CC FT METAL 96 96 Zinc (By similarity).
CC SQ SEQUENCE 340 AA; 36550 MW; 8DB47E14F8B8E9F7 CRC64;

Query Match 68.4%; Score 1210.5; DB 1; Length 340;
Best Local Similarity 69.0%; Pred. No. 5.1e-77;
Matches 236; Conservative 39; Mismatches 64; Indels 3; Gaps 1;

QY 1 MPIATPEVYAEMLGQAKQNSYAFPAINCTSETVNAAIKGFADAGSDGIIFSTGGAEFG 60
DB 1 MPIATPEVYNEMLDRAKAGKAFYPAINTVTSQTALRALRFAEASDGIQISTGAEFL 60

QY 61 SGLGVKDMVTGAVALAEETHVIAAKYPNVVALHTDHCPCDKLDSVVRPLLAI SAQRVSKG 120
DB 61 GQYSKDMVTGAVALAEFAHIAEKYPNVIALHTDHCPCDKLDSVVRPLLALSKRVEAG 120

QY 121 GNPLFQSHMWDGSAVPIDENLAIQAELLKAAAKIILEIEIGVVGGEDGVANEINEKL 180
DB 121 LGPLFQSHMWDGSAEPLADNLAIQAELLETRAAQIILEVEITPTGGEDGVSHINDSL 180

QY 181 YTSPEDEKTIETALGAGEHGKYLAAATFGNVHGVKPGNVKLRPDLIAQGOQVAAAKLGL 240
DB 181 YTTVDDAIRTAELGLGEGKRYLLAASFGNVHGVKPGNVKLRPDLIAQGOQVAAAKLGL 240

QY 241 PADAKPFDVFHGGSGSLKSEIEEALRYGVVKMNVDTDTQYAFTRPIAGHMTNVDGVLK 300
DB 240 --KESPFDFVFHGGSGSEIEIRTALENGVVKMNLDTDTQYAFTRPVAGHMFAYVDGVLK 297

QY 301 VDGEVGVKKVYDPRSYLKKAASMSQRVVQACNDLHCAGKSL 342
DB 298 VDGEVGNKKTYPDRTWGLAEAGMAARVVEATQHLRSAGNKI 339

RESULT 10
Q82E00

```

```
ID AC Q82EU0 PRELIMINARY; PRT; 340 AA.
DB 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE Putative fructose 1,6-bisphosphate aldolase.
GN Name-fba, OrderedLocusNames=SAV4523;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211431198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
CC -!- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycerone
CC phosphate + D-glyceraldehyde 3-phosphate.
CC -!- COFACTOR: Zinc (By similarity).
CC -!- PATHWAY: Glycolysis; sixth step.
CC -!- SIMILARITY: Belongs to the class II fructose-bisphosphate aldolase
CC family.
DR EMBL; AF005039; BAC72235.1; -.
DR HSSP; P11604; 1ZEN.
DR GO; GO:0004332; F:fructose-bisphosphate aldolase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006411; Fruct_bisp_bact.
DR InterPro; IPR000771; Kbp_aldolase.
DR Pfam; PF01116; F_bp_aldolase; 1.
DR ProDom; PD002376; K_bp_aldolase; 1.
DR TIGRFAMs; TIGR00167; cbbA; 1.
DR TIGRFAMs; TIGR01520; FruBisaldo II A; 1.
DR PROSITE; PS00602; ALDOLASE_CLASS_II_1; 1.
DR Complete proteome; Glycolysis; Lyase.
KW SEQUENCE 340 AA; 36817 MW; 996C777DCCF93DC6 CRC64;
SQ
Query Match 68.4%; Score 1210.5; DB 2; Length 340;
Best Local Similarity 69.0%; Pred. No. 5.1e-77;
Matches 236; Conservative 39; Mismatches 64; Indels 3; Gaps 1;

QY 1 MPIATPEVYAEMLGQAKNSYAPPAINTSSTVNAAIKGFADAGSDGIIQFSTGGAERG 60
DB 1 MPIATPEVYNEMLDRAKAGKAFYPAINTVSTQTLHAALGFAPAEASDGIQVISTGGAERL 60
QY 61 SGLGVKDMVTGAVALAEFTHTVIAKYPVNVALHTDHCPCDKLDVYRPLLAISAQRVSKG 120
DB 61 GGOYKDMVTGAVALAEFAHVAEKYPVTVALHTDHCPCDKLDVYRPLLAISSEERVKAG 120
QY 121 GNPLFOSHMDGSAVPIDENLATAQELLKAAAKIILEIEIGVCGEEDGVANEINEKL 180
DB 121 RNPLFOSHMDGSAETLADNLATAQELLEKARAIIILEIETPTGGEEDGVTHEINDSL 180
QY 181 YTSPEDEFTIEALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDLIAQQQVAAAKLGL 240
DB 181 YTTVEDAIRTAELGALGEGRYLLAASFGNVHGVYKPGNVVLRPELLKELNEGVAAKFG- 239
QY 241 PADAKPFDVFVHGGSGSLKSEIEEALRYGVVKNVNDTDTQYAFTRPIAGHMFTNYDGVLK 300
```

```
DB 240 --KADPFDFVFHGGSGSTEAEIRTALENGVWVKNVNDTDTQYAFTRPIAGHMFTNYDGVLK 297
QY 301 VDGEVGVKKVYDPRSYLKKAEASMSQVVOACNDLHCACKSL 342
DB 298 VDGEVGDKKTYDPTRWGKLAEASMAKRVTEACANLRSTGTGI 339
RESULT 11
Q6A687
ID Q6A687 PRELIMINARY; PRT; 341 AA.
AC Q6A687;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Fructose-bisphosphate aldolase (EC 4.1.2.13).
GN OrderedLocusNames=PPA2011;
OS Propionibacterium acnes.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Propionibacterineae; Propionibacteriaceae; Propionibacterium.
OX NCBI_TaxID=1747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KPA171202 / DSM 16379;
RX PubMed=15286373; DOI=10.1126/science.1100330;
RA Brueggemann H., Henne A., Hoster F., Liesegang H., Wierer A.,
RA Strittmatter A., Hujer S., Duerre P., Gottschalk G.;
RT "The complete genome sequence of Propionibacterium acnes, a commensal
RT of human skin.";
RL Science 305:671-673(2004).
CC -!- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycerone
CC phosphate + D-glyceraldehyde 3-phosphate.
CC -!- COFACTOR: Zinc (By similarity).
CC -!- PATHWAY: Glycolysis; sixth step.
CC -!- SIMILARITY: Belongs to the class II fructose-bisphosphate aldolase
CC family.
DR EMBL; AE017283; AAT83726.1; -.
DR GO; GO:0004332; F:fructose-bisphosphate aldolase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006411; Fruct_bisp_bact.
DR InterPro; IPR000771; Kbp_aldolase.
DR Pfam; PF01116; F_bp_aldolase; 1.
DR ProDom; PD002376; K_bp_aldolase; 1.
DR TIGRFAMs; TIGR00167; cbbA; 1.
DR TIGRFAMs; TIGR01520; FruBisaldo II A; 1.
DR PROSITE; PS00602; ALDOLASE_CLASS_II_1; 1.
DR PROSITE; PS00806; ALDOLASE_CLASS_II_2; 1.
DR Complete proteome; Glycolysis; Lyase.
KW SEQUENCE 341 AA; 36842 MW; F9B31CE28F9BA13F CRC64;
SQ
Query Match 67.6%; Score 1195.5; DB 2; Length 341;
Best Local Similarity 65.3%; Pred. No. 5.8e-76;
Matches 224; Conservative 45; Mismatches 71; Indels 3; Gaps 1;

QY 1 MPIATPEVYAEMLGQAKNSYAPPAINTSSTVNAAIKGFADAGSDGIIQFSTGGAERG 60
DB 1 MPIATPEVYEMIDRAKEQGFAYPAINTVSTQTLNAALOGFTGAGSDGIVQISTGGAEYA 60
QY 61 SGLGVKDMVTGAVALAEFTHTVIAKYPVNVALHTDHCPCDKLDVYRPLLAISAQRVSKG 120
DB 61 SQQTVKDMVTGAVALAEYAHVVAKNYPINVALHTDHCPCDKLDVYRPLLAISQERVDKG 120
QY 121 GNPLFOSHMDGSAVPIDENLATAQELLKAAAKIILEIEIGVCGEEDGVANEINEKL 180
DB 121 QDPLFNHMDGSALEVEENLQIABELLSRTAKAILIEIEVGAVGGEEDGVTEINEKL 180
QY 181 YTSPEDEFTIEALGAGEHGKYLAAATFGNVHGVYKPGNVKLRPDLIAQQQVAAAKLGL 240
DB 181 YTTVADGMTLEVLGEGERYTTALTFGNVHGVYKPGVYKLRPGLKIEIQDECCKYK- 239
QY 241 PADAKPFDVFVHGGSGSLKSEIEEALRYGVVKNVNDTDTQYAFTRPIAGHMFTNYDGVLK 300
```

```
Db 240 --KEKPFDLVPHGSGSTAQEIADAVSYGVKMNIDTDTQAFSPVVEHFHFKYDGMJK 297
QY 301 VDGEVGVKKVYDPRSYLKKASMSQRVQVQACNDLHCAGKSLT 343
Db 298 IDGEVGNKMYDPRSGWKKAGMAARIVEACEQLGSKGTSVS 340

RESULT 12
Q8G6T8 PRELIMINARY; PRT; 355 AA.
ID Q8G6T8 AC Q8G6T8
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Fructose-bisphosphate aldolase.
GN Name=fba; OrderedLocusNames=BL0550;
OS Bifidobacterium longum.
OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=216816;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 2705;
RX MEDLINE=22294977; PubMed=12381787; DOI=10.1073/pnas.212527599;
RA Schell M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,
RA Pessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M.,
RA Primrose R.D., Arigoni F.;
RT "The genome sequence of Bifidobacterium longum reflects its adaptation
RT to the human gastrointestinal tract.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427 (2002).
CC -!- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycero-
CC phosphate + D-glyceraldehyde 3-phosphate.
CC -!- COFACTOR: Zinc (By similarity).
CC -!- PATHWAY: Glycolysis; sixth step.
CC -!- SIMILARITY: Belongs to the class II fructose-bisphosphate aldolase
CC family.
DR EMBL; AE014676; AAN24374.1; -.
DR HSSP; P11604; 1B57.
DR GO; GO:0004332; F:fructose-bisphosphate aldolase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006411; Fruct_bisp_bact.
DR InterPro; IPR000771; Kbp_aldolase.
DR Pfam; PF01116; F_bp_aldolase; 1.
DR ProDom; PD002376; Kbp_aldolase; 1.
DR TIGRFAMs; TIGR00167; cdba; 1.
DR TIGRFAMs; TIGR01520; FruBisAldo II A; 1.
DR PROSITE; PS00602; ALDOLASE_CLASS_II_1; UNKNOWN_1.
DR PROSITE; PS00806; ALDOLASE_CLASS_II_2; 1.
KW Complete proteome; Glycolysis; Lyase.
SQ SEQUENCE 355 AA; 38349 MW; 96BE5FD6793B8EB CRC64;

Query Match 62.2%; Score 1101; DB 2; Length 355;
Best Local Similarity 61.9%; Pred. No. 2.7e-69;
Matches 219; Conservative 42; Mismatches 81; Indels 12; Gaps 2;
QY 1 MPIATPEVYAEMLGQAKNSYAPPAINTSSTSETVNAAIKGFADAGSDGI IQFSTGGAEFG 60
Db 1 MPIATPEVYAEMLGQAKNSYAPPAINTSSTSETVNAAIKGFADAGSDGI IQFSTGGAEFG 60
QY 61 SGLGVKDMVTGAVALAEFTHTVIAKYP-VNVALHTDHCPCDKLSDYVRLPLAISQRVSK 119
Db 61 SGRGVNDRVTGSLAAFAAHEVAARYPNITIALHTDHCAPQVLDWARPLLAHEADQVAR 120
QY 120 GGNPLFQSHMWGDSAPVIDENLAIQAELLKAAAAAKIILEIGVVGGEEDGVANEINEK 179
Db 121 GGEPTFQSHMWGDSVPLKENLDIAEELLDKSKAHTVLEIEIGAVGGEEDGHSANEIK 180
QY 180 LYTSPEDFKTIEALGAGHGKYLAAATFGNVGHVYKPGNVKLRPDL-----A 228
Db 181 LYSTPEDGLEVARRLGLGRGYMAAFTFGNVGHGAYKPGVVKLRPSLLGDIQARVARA 240
```

```
QY 229 QGOQVAAAKGLPADAKPFDVFPHGSGSLKSEIEALRYGVVVMNVDTDQVAFTRPIA 288
Db 241 EGGLEPSAAGIVDFPNGKPFELVFHGGSGRPEIAEAVSYGVKMNIDTDTQVAFTRPIA 300
QY 289 GHMFTNYDGLVKVDGEVGVKKVYDPRSYLKKASMSQRVQVQACNDLHCAGKSL 342
Db 301 DHVFENYDKVLKIDGEVGEKKFYDPRSGWKKAGDSMSARVVEACRQLGSAGKAL 354

RESULT 13
Q6ADV3 PRELIMINARY; PRT; 341 AA.
ID Q6ADV3 AC Q6ADV3
DT 25-OCT-2004 (Tremblrel. 28, Created)
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
DE Fructose 1,6-bisphosphate aldolase.
GN Name=fba; OrderedLocusNames=Lxx16680;
OS Leifsonia xyli (subsp. xyli).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococciaceae; Microbacteriaceae; Leifsonia.
OX NCBI_TaxID=59736;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CTCB07;
RX PubMed=15305603;
RA Monteiro-Vitorello C.B., Camargo L.E.A., Van Sluys M.A.,
RA Kitajima J.P., Truffi D., do Amaral A.M., Hatakava R.,
RA de Oliveira J.C.F., Wood D., de Oliveira M.C., Miyaki C.Y.,
RA Takita M.A., da Silva A.C.R., Furlan L.R., Carraro D.M., Camarotte G.,
RA Almeida N.F. Jr., Carter H., Coutinho L.L., El-Dorri H.A.,
RA Ferro M.T., Gagliardi P.R., Gigliotti E., Goldman M.H.S.,
RA Goldman G.H., Kimura E.T., Ferro E.S., Kuramae E.E., Lenos E.G.M.,
RA Lenos M.V.F., Mauro S.M.Z., Machado M.A., Marino C.L., Menck C.F.,
RA Nunes L.R., Oliveira R.C., Pereira G.G., Siqueira W., de Souza A.A.,
RA Tsai S.M., Zanca A.S., Simpson A.J.G., Brumbley S.M., Setubal J.C.;
RT "The genome sequence of the Gram-positive sugarcane pathogen Leifsonia
RT xyli subsp. xyli.";
RL MOL. Plant Microbe Interact. 17:827-836 (2004).
CC -!- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycuro-
CC phosphate + D-glyceraldehyde 3-phosphate.
CC -!- COFACTOR: Zinc (By similarity).
CC -!- PATHWAY: Glycolysis; sixth step.
CC -!- SIMILARITY: Belongs to the class II fructose-bisphosphate aldolase
CC family.
DR EMBL; AE016822; AAT89443.1; -.
DR GO; GO:0004332; F:fructose-bisphosphate aldolase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006411; Fruct_bisp_bact.
DR InterPro; IPR000771; Kbp_aldolase.
DR Pfam; PF01116; F_bp_aldolase; 1.
DR ProDom; PD002376; Kbp_aldolase; 1.
DR TIGRFAMs; TIGR00167; cdba; 1.
DR TIGRFAMs; TIGR01520; FruBisAldo II A; 1.
DR PROSITE; PS00602; ALDOLASE_CLASS_II_1; 1.
DR PROSITE; PS00806; ALDOLASE_CLASS_II_2; 1.
KW Complete proteome; Glycolysis; Lyase.
SQ SEQUENCE 341 AA; 36226 MW; 4D7FD48E37227C8D CRC64;

Query Match 61.6%; Score 1089; DB 2; Length 341;
Best Local Similarity 61.0%; Pred. No. 1.8e-68;
Matches 208; Conservative 51; Mismatches 80; Indels 2; Gaps 1;
QY 1 MPIATPEVYAEMLGQAKNSYAPPAINTSSTSETVNAAIKGFADAGSDGI IQFSTGGAEFG 60
Db 1 MPIATPEVYAEMLGQAKNSYAPPAINTSSTSETVNAAIKGFADAGSDGI IQFSTGGAEFG 60
QY 61 SGLGVKDMVTGAVALAEFTHTVIAKYPVNVALHTDHCPCDKLSDYVRLPLAISQRVSK 120
Db 61 AGQSVKARATGALAFARFATEVAKNPITVALHTDHCPCDKLSDYVRLPLAISQRVSK 120
QY 121 GNPLFQSHMWGDSAPVIDENLAIQAELLKAAAAAKIILEIGVVGGEEDGVANEINEK 180
```


Db	121	RPFIQSHMDGSAVPLADNLEIAKQMTERTRAINAILLEVEITGVGVGGEDGVSHEINHL	180
Qy	181	YTSPEDEFTKEALCAGEHGKYLAAATFGNVHGVYKPGNVKLRDPDILAQGOQVAAKLGL	240
Db	181	YTTLDDATQTVEALGLGNGRYMAALTFGNVHGVYKPGNVKLRPELLKXIDPGLAAKYG-	239
Qy	241	PADAKPDPFVPHGGSGSLKSRIEEALRYGVYKMNVDTDQYATRPPIAGHMFTNYDGLVK	300
Db	240	-TSPKPLDLVPHGGSGSTDEIEAVRNGVYKMNITDQYATRSIAGYMGFNVDGLVK	298
Qy	301	VDGEVGKVKVDPRSYLKAQASMSQRVYQACNDLHCAGKS	341
Db	299	IDGSEGNKKVDPRAGWKVAESAARVVEATQQLGSAGHS	339

```

RESULT 14
Q9CJX8
ID Q9CJX8 PRELIMINARY; PRT; 359 AA.
AC Q9CJX8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
Fba.
Name=fba; OrderedLocusNames=PM1861;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
NCBI_TaxID=747;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=PM70;
RC MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
RX May B.J., Zhang Q., Li L.L., Faustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycerone
CC phosphate + D-glyceraldehyde 3-phosphate.
CC -1- COPACTOR: Zinc (By similarity)..
CC -1- PATHWAY: GLYCOLYSIS; sixth step.
CC -1- SIMILARITY: Belongs to the class II fructose-bisphosphate aldolase
CC family.
CC EMBL; AE006224; AAK03945.1; -.
DR HSP; P11604; 1B57.
DR GO; GO:0004332; F:fructose-bisphosphate aldolase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008270; P:zinc ion binding; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001194; ATPase_a/bcentre.
DR InterPro; IPR006411; Fruct_bisp_bact.
DR InterPro; IPR000771; K_bp_aldolase.
DR Pfam; PF01116; F_bp_aldolase; 1.
DR ProDom; PD002376; K_bp_aldolase; 1.
DR TIGRfams; TIGR00167; cbba; 1.
DR TIGRfams; TIGR01520; FruBisAlco_II_A; 1.
DR PROSITE; PS00602; ALDOLASE_CLASS_II_1; 1.
DR PROSITE; PS00806; ALDOLASE_CLASS_II_2; 1.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
KW Complete proteome; Glycolysis; Lyase.
SQ SEQUENCE 359 AA; 39247 MW; 6c7409B40EA6260A CRC64;

```

	Query Match	40.1%	Score 708.5;	DB 2;	Length 359;
	Best Local Similarity	44.0%;	Pred. No. 1.1e-41;		
	Matches 158;	Conservative 58;	Mismatches 104;	Indels 39;	Gaps 9;
Qy	3	IATPEVVAEMIGQAKQNSYAPPAINCTSSETVNAAIKGFADAGSDGIIIQSTTGGAEFGSG	62		
Db	12	VVTGDVDQKVFAYAKANNFAIPNCVCGSDSVNAVLETAARVKAPVIIQFNSNGCAQFPVAG	71		
Qy	63	LGVIKDM-----VTGAVALAEFTHVIAAKYPVVNALHTDHCPOKLDSYVVPILAIASQR	116		
Db	72	KGLKPASGARTDVLTGATAGAKHVALAEYGVPIIITDTHAAK-KLLPWIIDGLEAGEEH	130		

		E	175
Qy	117	VSKGPNLPFOSHMMWDGSAVPIDENLAIQAELLKAAAANKIILEITBIVGVGBEDGVAN-E	175
Db	131	FAETGKLPLFSHHMIDLSEEPMEENNAICREYLARMDMKMGMTLEIBIGITGGBEDGVNSD	190
Qy	176	INB-KLVTSPEDFEKTTEALGAGEHGKYLLAATTCGNVHGVVKPGNVKLRPDLTLAGOQOVA	234
Db	191	VEBSKLYIQEDVLVYVDQLNP-VSPRPTVAAGFNHVGVKPGNVKLRPSILGASQEFV	2419
Qy	235	AALKGLPADAKPFDPFHGGSGSLKSEIEEALRYGVVRMNVDTDQTAPTRPIAGHMFTN	294
Db	250	SKERGLP--AKSIDVFPHGGSGSSREEIREALSYGAIKMNIDTDTQWA-----S	296
Qy	295	YDGVLK-----VDGEVG-----KKVYDPRSVLYKKAESNSQRVQACNDLHC	337
Db	297	WEGTLKFYNANODLYOLOGLNPEGPDAPNKYYDPVRLRWKEEMSMKRLEOSFDCLNC	355

```

RESULT 15
Q7VMV1
ID Q7VMV1 PRELIMINARY; PRT; 358 AA.
AC Q7VMV1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Fructose-bisphosphate aldolase class II.
GN Name=fbaA; OrderedLocusNames=HD0864;
OS Haemophilus ducreyi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
NCBI_TaxID=730;
[1]
SEQUENCE FROM N.A.
SFRAIN=35000HP / ATCC 700724;
RA Johnson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,
RA Munson L., Nguyen D., Wang J., Forst C., Hood L.;
RT "The complete genome sequence of Haemophilus ducreyi."
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC -! CATALYTIC ACTIVITY: D-fructose 1,6-bisphosphate = glycerone
CC phosphate + D-glyceraldehyde 3-phosphate.
CC -! COFACTOR: Zinc (By similarity)..
CC -! PATHWAY: Glycolysis; sixth stage.
CC -! SIMILARITY: Belongs to the class II fructose-bisphosphate aldolase
CC family.
EMBL; AE017153; AAP95752.1; -.
HSP; P11604; LGVN.
DR GO; GO:0004332; F:fructose-bisphosphate aldolase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000194; ATPase_a/bcentre.
DR InterPro; IPR006411; Fruct_Bisp_bact.
DR InterPro; IPR000771; K_bp_aldolase.
DR Pfam; PF0116; F_bp_aldolase; 1.
DR ProDom; PD002376; K_bp_aldolase; 1.
DR TIGRfams; TIGR00167; cbba; 1.
DR TIGRfams; TIGR01520; FruBisAldo II A; 1.
DR PROSITE; PS00602; ALDOLASE_CLASS_II_1; 1.
DR PROSITE; PS00806; ALDOLASE_CLASS_II_2; 1.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
KW Complete proteome; Glycolysis; Lyase.
SQ SEQUENCE 358 AA; 39248 MW; 171D601E27F54B9D CRC64;

```

[illegible]

QY 117 VSKGNPLFOSHMDGSAVPIDENLAIQAELLKAAAAAKIILEIEIGVVGGEEGCVAN-E 175
Db 130 FAETGKPLFSSHMLDSEEPMEENMAICREYLARMSKMDMTLEIEIGITGGEEGVDNSN 189
QY 176 INE-KLYTSPEDFEKTIIBALGAGEH---GKYLAAATFGNVHGVYKPGNVKLRPDILAOGQ 231
Db 190 VDESRLYTQPEDVLVYDQL----HPISPRFTIAAFGCVHGVYKPGNVKLRKPSILGDSQ 245
QY 232 QVAAAKLGLPADAKPFDVFHGGGSLKSEIEEALRYGVVKNVNDTDTQYAFTRPIAGHM 291
Db 246 EFVSAQRNLP--AKSINFVHGGGSRREIRQAISYGAIKMNIDTDTQWA----- 294
QY 292 FTNYDGLVK-----VDGEVG-----VKVYDPRSVLKKAAASMSORVVOACNDL 335
Db 295 --SWDGILNFYKANEAYLQGQGNPDGSPAPNKKYIDPRVWLKRMESMSKRLSQSFEDL 352
QY 336 HC 337
Db 353 NC 354

Search completed: October 15, 2005, 04:35:51
Job time : 79.0538 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:20:35 ; Search time 30.0403 Seconds
(without alignments)
1454.845 Million cell updates/sec

Title: US-10-617-038-5
Perfect score: 595
Sequence: 1 MGEHAIKRHRMRQKTKHPL.....TTESSFEHSRSPSQSDLL 113

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	595	100.0	113	8	ADI37284
2	78.5	13.2	188	4	ABb66418 M. tuberc
3	75	12.6	539	7	ABO80809 Drosophil
4	73.5	12.4	664	5	ABP73358 Pseudomon
5	73.5	12.4	1232	5	ABP73358 Candida a
6	73.5	12.4	1244	8	ABP52101 Homo sapi
7	73.5	12.4	1251	8	ADh22536 Human tra
8	73.5	12.4	1259	8	ADh22533 Human tra
9	73.5	12.4	1260	8	ABm84850 Human dia
10	73.5	12.4	1286	8	ABm83184 Human dia
11	73.5	12.4	1294	8	ADP56345 Human PRO
12	73.5	12.4	1295	8	ABm83182 Human dia
13	72.5	12.2	1527	5	ABm83183 Human dia
14	71	11.9	809	6	AuA80055 Leuconost
15	71	11.9	1125	4	ADA20755 Soybean p
16	70	11.8	561	4	ABb69365 Drosophil
17	70	11.8	561	6	Aau36508 Pseudomon
18	70	11.8	597	7	ABU38909 Protein e
19	70	11.8	653	8	ABO68393 Pseudomon
20	70	11.8	751	8	ADs85117 Mouse ato
21	69.5	11.7	1527	7	ABR82253 Human nuc
22	68.5	11.5	286	4	ADc54807 Leuconost
23	68.5	11.5	506	6	AAG31033 C glutami
24	68.5	11.5	2920	8	ABu33977 Protein e
25	68	11.4	326	3	ADo29246 Mouse GPC
					AAb18328 Plasmodiu

26	68	11.4	1430	2	AAY31949
27	67.5	11.3	433	3	AAY59355
28	67.5	11.3	433	3	AAY59355
29	67	11.3	188	7	ABO73047 Pseudomon
30	67	11.3	1590	7	ADb93657 Streptoco
31	66.5	11.2	278	7	ABO79966 Pseudomon
32	66.5	11.2	401	7	ABO70528
33	66.5	11.2	565	6	ABU20710
34	66	11.1	130	5	ABP62181
35	66	11.1	130	5	ABP62183
36	66	11.1	247	7	ABO78673
37	66	11.1	840	4	ABBS2606
38	66	11.1	1024	5	ABBO4818
39	66	11.1	1346	6	AAW31346
40	66	11.1	1346	6	ABG73371
41	66	11.1	2107	7	ABM85870
42	66	11.1	2541	3	AAb41087
43	66	11.1	2541	5	ABBB1459
44	66	11.1	2541	5	ADG90450
45	66	11.1	2541	6	ABR47614

ALIGNMENTS

RESULT 1
ADI37284
ID ADI37284 standard; protein; 113 AA.
AC ADI37284;
DT 22-APR-2004 (first entry)
DE M. tuberculosis low oxygen induced antigen Rv0572c SEQ ID NO:5.
KW mycobacterial infection; vaccine; tuberculosis;
KW Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;
KW low oxygen induced antigen.
OS Mycobacterium tuberculosis.
XX WO2004006952-A2.
XX 22-JAN-2004.
XX 08-JUL-2003; 2003WO-DK000477.
XX 13-JUL-2002; 2002DK-00001098.
(STAT-) STATENS SERUM INST.
Andersen P, Rosenkrands I, Stryhn A;
WPI; 2004-122778/12.
N-PSDB; ADI37329.
Use of one or more polypeptides or their fragments, which are expressed during the latent stage of the mycobacterial infection, and/or nucleic acids encoding the polypeptides, for a therapeutic vaccine against tuberculosis.
Claim 3; SEQ ID NO 5; 76pp; English.
The present invention describes polypeptides or their fragments, which are expressed during the latent stage of a mycobacterial infection, and/or nucleic acids encoding the polypeptides, which are useful for creating a therapeutic vaccine against tuberculosis. Also described: (1) a therapeutic vaccine against tuberculosis comprising one or more polypeptides; (2) a method for treating an animal, including a human being, with tuberculosis caused by virulent mycobacteria, e.g. by Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for immunising an animal, including a human being, against tuberculosis caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis

CC caused by virulent mycobacteria in an animal, including a human being;
CC (5) a method for diagnosing previous or ongoing infection with a virulent
CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
CC infection in a subject. The polypeptides have antibacterial activities,
CC and can be used in vaccines and in gene therapy. The polypeptides are
CC useful for the manufacture of a therapeutic vaccine for treating an
CC individual who is infected by a virulent mycobacterium, e.g. M.
CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
CC The present sequence represents a low oxygen induced antigen, which is
CC used in the exemplification of the present invention.

XX SQ Sequence 113 AA;
Query Match 100.0%; Score 595; DB 8; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.9e-65;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGEHAIKRMQRKPTKHPKLAQKRGARILVFTDDPRRSVLIVPGCHLDSMRREKNAYVFQ 60
DB 1 MGEHAIKRMQRKPTKHPKLAQKRGARILVFTDDPRRSVLIVPGCHLDSMRREKNAYVFQ 60
QY 61 DGNALVGMVSGGTVEYDADRTYVQLTDGRHTTSSFEHSSPSRSPQSDDL 113
DB 61 DGNALVGMVSGGTVEYDADRTYVQLTDGRHTTSSFEHSSPSRSPQSDDL 113

RESULT 2
ABB66418
ID ABB66418 standard; protein; 188 AA.
XX AC ABB66418;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 26046.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX DR N-PSDB; ABL10521.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX Disclosure; SEQ ID NO 26046; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-
XX ABBS72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 188 AA;
Query Match 13.2%; Score 78.5; DB 4; Length 188;
Best Local Similarity 30.4%; Pred. No. 0.48;
Matches 34; Conservative 14; Mismatches 45; Indels 19; Gaps 6;
QY 1 MGEHAIKRMQRKPTKHPKLAQKRGARILVFTDDPRRSVLIVPGCHLDSMRREKNAY 57
DB 43 MFAHKPSTHKRQYLTTNPDPSPSA-TFDFDRDRSVLEIP----DRSYRLVLRLAY 97
QY 58 YFQDGNALVGMVSGGTVEYDADD-----RTYVQL-----TDGRHTTSSFE 100
DB 98 GLLDSGASISCI--GGOLEKAATFENFKSLTIELLTKTSFGRSSSDSFE 147
RESULT 3
ABO80809
ID ABO80809 standard; protein; 539 AA.
XX AC ABO80809;
XX DT 29-JUL-2004 (first entry)
XX DE Pseudomonas aeruginosa polypeptide #12984.
XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX OS Pseudomonas aeruginosa.
XX PN US6551795-B1.
XX PD 22-APR-2003.
XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
XX DR N-PSDB; ABD14380.
XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, of
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX Disclosure; SEQ ID NO 29555; 455pp; English.
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences ABO67826-
XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html
XX SQ Sequence 539 AA;
Query Match 12.6%; Score 75; DB 7; Length 539;
Best Local Similarity 32.5%; Pred. No. 5.4;

Matches 25; Conservative 13; Mismatches 31; Indels 8; Gaps 4;	
QY 3	EHAKEHMRQKPTKHLAQKRGARIL-VFTDDPRRSVLIVPGCHLDSMRREKNAYFQ- 60
Db	44 QRAVREYRVQPPAERLLRLSLGRLGAKPQR-----VPSCHPDSRSDQPNNAEVNF 99
QY 61	DGNALVGMVMSGTVBY 77
Db	100 DG--LVGPHTNYGGLSY 114
RESULT 4	
ABP73358	
ID	ABP73358 standard; protein; 664 AA.
AC	ABP73358;
XX	
XX	30-JAN-2003 (first entry)
DT	
DE	Candida albicans essential protein SEQ ID NO 7195.
XX	
KW	Fungus; Yeast; tetracyclin; promoter; GRACE strain; biosynthesis;
KW	signal transduction; DNA replication; cell division; growth;
KW	proliferation; Candida albicans; fungicide; antifungal.
XX	
OS	Candida albicans.
XX	
PN	WO200253728-A2.
XX	
PD	11-JUL-2002.
XX	
PF	26-DEC-2001; 2001WO-US049486.
XX	
PR	29-DEC-2000; 2000US-0259128P.
PR	20-FEB-2001; 2001US-00792024.
PR	22-AUG-2001; 2001US-0314050P.
XX	
PA	(ELIT-) ELITRA PHARM INC.
XX	
PI	Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;
XX	
DR	WPI; 2002-566694/60.
DR	N-PSDB; AB231908.
XX	
PT	Constructing strains for identifying gene products as effective targets
PT	for therapeutic intervention, by inactivating in the strain one allele of
PT	a gene and placing other allele of the gene under conditional expression.
XX	
PS	Claim 44; SEQ ID NO 7195; 167pp + Sequence Listing; English.
XX	
CC	The invention relates to constructing (M1) a strain of diploid fungal
CC	cells in which both alleles of a gene are modified, comprising modifying
CC	one allele by insertion or replacement by a cassette having an
CC	expressible selectable marker and modifying other allele by
CC	recombination, of a promoter replacement fragment with a heterologous
CC	promoter, so that expression of the second allele is regulated by the
CC	cells in which both alleles of a gene are modified. The diploid fungal
CC	cells having both alleles modified are useful for identifying a gene that
CC	is essential to the survival or growth of a fungus, a gene that
CC	contributes to the virulence and/or pathogenicity of a fungus, a gene
CC	that contributes to the resistance of a diploid fungus to an antifungal
CC	agent, an antifungal agent that inhibits the growth of a diploid fungus
CC	and for identifying a therapeutic agent for treatment of a mammalian
CC	disease. (M1) is useful for identifying a compound which modulates the
CC	activity of a gene product, preferably enzymatic activity, carbon
CC	compound catabolism, biosynthetic, transporter, transcriptional,
CC	translational, signal transduction, DNA replication and cell division
CC	activity. The method is useful for identifying a compound having the
CC	ability to inhibit growth or proliferation of C. albicans cells and for
CC	treating infection by C. albicans. The present sequence is that of an
CC	essential Candida albicans protein used in the method of the invention.
CC	Note: The sequence data for this patent is not represented in the printed
specification but is based on sequence information supplied to Derwent by	
the European Patent Office	
XX	
SQ	Sequence 664 AA;
Query Match 12.4%; Score 73.5; DB 5; Length 664;	
Best Local Similarity 24.1%; Pred. No. 11;	
Matches 27; Conservative 21; Mismatches 49; Indels 15; Gaps 4;	
QY 12	QRKPTKHLAQKRGARILVFTDDPRRSVLIVPGCHLDSMRREKNAYFQDQ---NALVG 67
Db	74 QTKLAKNSTPVSCKTKRDTIGTSNRKQRTVDNHHLTLPOLKKRK---FEDNTPEKTPAG 130
QY 68	MVMSGTVEYDADRTYVVLTDGRHTTESSEHSPSRSS-----PQSDDL 113
Db	131 IPVSSPTTDYETANNTFVGLTTNGTPTTTTA--KTPSNSLLETLPQIDEI 180
RESULT 5	
ABP52101	
ID	ABP52101 standard; protein; 1232 AA.
XX	
AC	ABP52101;
XX	
DT	10-OCT-2002 (first entry)
XX	
DE	Homo sapiens ABC transporter ABCB4 protein SEQ ID NO:53.
XX	
KW	ATP-binding cassette transporter; ABC transporter; modulation; D loop;
KW	cancer; bacterial infection; fungal infection; protozoal infection;
KW	antibacterial; fungicide; protozoacide.
XX	
OS	Homo sapiens.
XX	
PN	EP1217066-A1.
XX	
PD	26-JUN-2002.
XX	
PF	21-DEC-2000; 2000EP-00870316.
XX	
PR	21-DEC-2000; 2000EP-00870316.
XX	
PA	(UYGE-) UNIV GENT.
XX	
DR	WPI; 2002-550404/59.
XX	
PT	Modulating activity of ATP-binding cassette (ABC) transporters by
PT	influencing dimerization of nucleotide binding domains through use of D
PT	loop sequence of an ABC transporter, or its antisense peptide or peptide
PT	mimetic.
XX	
PS	Disclosure; Fig 3; 290pp; English.
XX	
CC	The present invention describes a method (M1) for modulating the activity
CC	of ATP-binding cassette (ABC) transporters by influencing the
CC	dimerization of the nucleotide binding domains comprising using: (a) a
CC	polypeptide (polyp) consisting of 5-50 amino acids comprising the D loop
CC	sequence of an ABC transporter (ABP52049 to ABP52091); (b) a polyp
CC	consisting of the D loop sequence of an ABC transporter; (c) a peptide
CC	mimetic or antisense peptide of (a) or (b). ABC transporters have
CC	antibacterial, fungicide and protozoacide activities. (M1) is useful for
CC	selectively modulating the activity of ABC transporters belonging to the
CC	group of multidrug transporter/P-glycoproteins. Bacterial, fungal or
CC	protozoal ABC transporters are involved in the infection of a mammal. (M1)
CC	in the induction of resistance to antibiotics or drugs in a mammal. (M1)
CC	is useful for preventing, treating or alleviating diseases associated
CC	with functionality of an ABC transporter. ABP52092 to ABP52140 represent
CC	ABC transporter proteins given in the exemplification of the present
CC	invention
XX	
SQ	Sequence 1232 AA;
Query Match 12.4%; Score 73.5; DB 5; Length 1232;	

Best Local Similarity 24.2%; Pred. No. 25;
Matches 23; Conservative 21; Mismatches 32; Indels 19; Gaps 5;
QY 18 HPLAQKRGARILVFTDDPRRSVLIVPGCHLDSMRREKNAYYPQDGNALVGMVWSGGTV-- 75
DB 589 HRLSTVRNADVIAGFED---GVIVEQSGSHSELMKKE--GVYFK-----LVNMQTSQSQIQS 639
QY 76 -EYDADRTYVVLQTDG-----RHTTESSFEHS 102
DB 640 BEFELNDEKAATRMAPNGWKSRLFRHSTQKNLKNS 674
RESULT 6
ID ADH22536
XX ADH22536 standard; protein; 1244 AA.
AC ADH22536;
DT 11-MAR-2004 (first entry)
DE Human transporter & ion channel (TRICH) protein SeqID34.
XX human; transporters and ion channel; TRICH; cell proliferative;
KW arteriosclerosis; cancer; autoimmune/inflammatory; AIDS; asthma;
KW neurological; epilepsy; stroke; developmental; Cushing's syndrome;
KW hypothyroidism; infection; gene therapy; cytostatic; antiinflammatory;
KW immunosuppressive; antiasthmatic; anticonvulsant; nootropic;
KW neuroprotective.
XX OS Homo sapiens.
XX PN WO2003093444-A2.
XX PD 13-NOV-2003.
XX PF 02-MAY-2003; 2003WO-US014026.
XX PR 03-MAY-2002; 2002US-0377435P.
XX PR 03-MAY-2002; 2002US-0377444P.
XX PR 05-JUN-2002; 2002US-0386497P.
XX PR 11-JUN-2002; 2002US-0388180P.
XX PA (INCY-) INCYTE CORP.
XX PI Baughn MR, Becha SD, Bulloch SA, Chang H, Elliott VS;
PI Emerling BM, Griffin JA, Hafalia AJA, Ison CH, Jackson AA, Jiang X;
PI Jin P, Kable AE, Khare R, Lee SY, Lee S, Mason PM, Marquis JP;
PI Ramkumar J, Richardson TW, Swarnakar A, Tran UK, Chawla NK;
PI Wilson AD;
XX DR WPI; 2004-022655/02.
XX DR N-PSDB; ADH22602.
XX PT New human transporters and ion channels (TRICH), useful for diagnosing,
PT treating and preventing diseases or conditions associated with the
PT aberrant TRICH expression e.g. cancer, AIDS, atherosclerosis, epilepsy,
PT or infections.
XX PS Claim 1; SEQ ID NO 34; 448pp; English.
XX CC This invention relates to novel isolated polynucleotides identified as
CC human transporters and ion channels (TRICH), and the encoded polypeptides
CC thereof. Specifically, it describes using these TRICH molecules, as well
CC as agonists, antagonists, antibodies, expression vectors and host cells,
CC in appropriate screening and toxicity assays to assess the effects of
CC exogenous compounds on TRICH expression. The present invention describes
CC TRICH compositions that are useful in the diagnosis, treatment and
CC prevention of various disorders such as cell proliferative (e.g.
CC arteriosclerosis, cancer), autoimmune/inflammatory (e.g. AIDS, asthma),
CC neurological (e.g. epilepsy, stroke) and developmental (e.g. Cushing's
CC syndrome hypothyroidism) and for infections. Accordingly, these TRICH
CC molecules can be used for gene therapy purposes and exhibit various
CC activities such as cytostatic, antiinflammatory, immunosuppressive,

CC antiasthmatic, anticonvulsant, nootropic and neuroprotective.
CC Furthermore, a microarray is useful in monitoring or measuring protein-
CC protein interactions, drug-target interactions and gene expression
CC profiles. This polypeptide sequence is a human TRICH protein of the
CC invention.
XX SQ Sequence 1244 AA;
Query Match 12.4%; Score 73.5; DB 8; Length 1244;
Best Local Similarity 24.2%; Pred. No. 25;
Matches 23; Conservative 21; Mismatches 32; Indels 19; Gaps 5;
QY 18 HPLAQKRGARILVFTDDPRRSVLIVPGCHLDSMRREKNAYYPQDGNALVGMVWSGGTV-- 75
DB 589 HRLSTVRNADVIAGFED---GVIVEQSGSHSELMKKE--GVYFK-----LVNMQTSQSQIQS 639
QY 76 -EYDADRTYVVLQTDG-----RHTTESSFEHS 102
DB 640 BEFELNDEKAATRMAPNGWKSRLFRHSTQKNLKNS 674
RESULT 7
ADH22533
ID ADH22533 standard; protein; 1251 AA.
XX ADH22533;
AC ADH22533;
DT 11-MAR-2004 (first entry)
XX Human transporter & ion channel (TRICH) protein SeqID31.
XX human; transporters and ion channel; TRICH; cell proliferative;
KW arteriosclerosis; cancer; autoimmune/inflammatory; AIDS; asthma;
KW neurological; epilepsy; stroke; developmental; Cushing's syndrome;
KW hypothyroidism; infection; gene therapy; cytostatic; antiinflammatory;
KW immunosuppressive; antiasthmatic; anticonvulsant; nootropic;
KW neuroprotective.
XX OS Homo sapiens.
XX XX WO2003093444-A2.
XX PN 13-NOV-2003.
XX PD 02-MAY-2003; 2003WO-US014026.
XX PF 03-MAY-2002; 2002US-0377435P.
XX PR 03-MAY-2002; 2002US-0377444P.
XX PR 05-JUN-2002; 2002US-0386497P.
XX PR 11-JUN-2002; 2002US-0388180P.
XX PA (INCY-) INCYTE CORP.
XX PI Baughn MR, Becha SD, Bulloch SA, Chang H, Elliott VS;
PI Emerling BM, Griffin JA, Hafalia AJA, Ison CH, Jackson AA, Jiang X;
PI Jin P, Kable AE, Khare R, Lee SY, Lee S, Mason PM, Marquis JP;
PI Ramkumar J, Richardson TW, Swarnakar A, Tran UK, Chawla NK;
PI Wilson AD;
XX DR WPI; 2004-022655/02.
XX DR N-PSDB; ADH22599.
XX PT New human transporters and ion channels (TRICH), useful for diagnosing,
PT treating and preventing diseases or conditions associated with the
PT aberrant TRICH expression e.g. cancer, AIDS, atherosclerosis, epilepsy,
PT or infections.
XX PS Claim 1; SEQ ID NO 31; 448pp; English.
XX CC This invention relates to novel isolated polynucleotides identified as
CC human transporters and ion channels (TRICH), and the encoded polypeptides
CC thereof. Specifically, it describes using these TRICH molecules, as well
CC as agonists, antagonists, antibodies, expression vectors and host cells,
CC in appropriate screening and toxicity assays to assess the effects of
CC exogenous compounds on TRICH expression. The present invention describes
CC TRICH compositions that are useful in the diagnosis, treatment and
CC prevention of various disorders such as cell proliferative (e.g.
CC arteriosclerosis, cancer), autoimmune/inflammatory (e.g. AIDS, asthma),
CC neurological (e.g. epilepsy, stroke) and developmental (e.g. Cushing's
CC syndrome hypothyroidism) and for infections. Accordingly, these TRICH
CC molecules can be used for gene therapy purposes and exhibit various
CC activities such as cytostatic, antiinflammatory, immunosuppressive,

CC in appropriate screening and toxicity assays to assess the effects of
CC exogenous compounds on TRICH expression. The present invention describes
CC TRICH compositions that are useful in the diagnosis, treatment and
CC prevention of various disorders such as cell proliferative (e.g.
CC arteriosclerosis, cancer), autoimmune/inflammatory (e.g. AIDS, asthma),
CC neurological (e.g. epilepsy, stroke) and developmental (e.g. Cushing's
CC syndrome hypothyroidism) and for infections. Accordingly, these TRICH
CC molecules can be used for gene therapy purposes and exhibit various
CC activities such as cytostatic, antiinflammatory, immunosuppressive,
CC antiaesthetic, anticonvulsant, nootropic and neuroprotective.
CC Furthermore, a microarray is useful in monitoring or measuring protein-
CC protein interactions, drug-target interactions and gene expression
CC profiles. this polypeptide sequence is a human TRICH protein of the
CC invention.
XX
SQ Sequence 1251 AA;

Query Match 12.4%; Score 73.5; DB 8; Length 1251;
Best Local Similarity 24.2%; Pred. No. 25;
Matches 23; Conservative 21; Mismatches 32; Indels 19; Gaps 5;
QY 18 HPLAQKRGARILVFTDPRRSVLIVPGCHLDSMRREKNAYIFQDGNALGMVVGTV-- 75
Db 589 HRLSTVRNADVIAGFED---GVIVEGSHSELMKKE--GVYFK---LVNMQTSGSQIQS 639
QY 76 -EYDADRTYVVLTDG-----RHTTSSFEHS 102
Db 640 EEFELNDEKAATRMAPNGWKSLFRHSTQKNLKNS 674

RESULT 8
ABM84850
ID ABM84850 standard; protein; 1259 AA.
AC ABM84850;
XX
XX
DT 18-NOV-2004 (first entry)
DE Human diagnostic and therapeutic pprotein SEQ ID NO:5099.
XX
XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.

OS Homo sapiens.
XX
PN WO2004023973-A2.
PD 25-MAR-2004.
XX
PF 12-SEP-2003; 2003WO-US028227.
XX
PR 12-SEP-2002; 2002US-0410259P.
PR 12-SEP-2002; 2002US-0410260P.
XX
PA (INCY-) INCYTE CORP.
XX
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtson ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patry S, Shi X, Suarez CJ;
XX
DR WPI: 2004-329368/30.
DR N-PSDB; ACN43502.

XX
PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.

PS Claim 27; Page; 190pp; English.

XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC polynucleotide and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorders,
CC disorder, neurological disorders, gastrointestinal disorders, endocrine
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ Sequence 1259 AA;

Query Match 12.4%; Score 73.5; DB 8; Length 1259;
Best Local Similarity 24.2%; Pred. No. 26;
Matches 23; Conservative 21; Mismatches 32; Indels 19; Gaps 5;
QY 18 HPLAQKRGARILVFTDPRRSVLIVPGCHLDSMRREKNAYIFQDGNALGMVVGTV-- 75
Db 589 HRLSTVRNADVIAGFED---GVIVEGSHSELMKKE--GVYFK---LVNMQTSGSQIQS 639
QY 76 -EYDADRTYVVLTDG-----RHTTSSFEHS 102
Db 640 EEFELNDEKAATRMAPNGWKSLFRHSTQKNLKNS 674

RESULT 9
ABM83184
ID ABM83184 standard; protein; 1260 AA.
AC ABM83184;
XX
XX
DT 18-NOV-2004 (first entry)
DE Human diagnostic and therapeutic pprotein SEQ ID NO:3433.
XX
XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
OS Homo sapiens.
XX
PN WO2004023973-A2.
XX
PD 25-MAR-2004.
XX
PF 12-SEP-2003; 2003WO-US028227.
XX
PR 12-SEP-2002; 2002US-0410259P.
PR 12-SEP-2002; 2002US-0410260P.
XX
PA (INCY-) INCYTE CORP.
XX
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtson ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patry S, Shi X, Suarez CJ;
XX
DR WPI: 2004-329368/30.
DR N-PSDB; ACN41836.

XX
PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or

PT in gene mapping.
 PS Claim 27; Page; 190pp; English.
 XX
 CC The invention relates to novel diagnostic and therapeutic polynucleotides
 CC selected from one of the 2722 sequences defined in the specification. A
 CC polynucleotide of the invention may have a use in gene therapy. The human
 CC diagnostic and therapeutic polynucleotides (dichp) or polypeptides may be
 CC used to diagnose a particular condition, disease or disorder associated
 CC with human molecules, e.g. cell proliferative disorders,
 CC autoimmune/inflammatory disorder, developmental disorder, endocrine
 CC disorder, neurological disorders, gastrointestinal disorders, or
 CC infections caused by virus, bacteria, fungi or parasite. The dichp
 CC molecules may also be used in genetic mapping, in identifying individuals
 CC from minute biological samples, in detecting single nucleotide
 CC polymorphisms, as molecular weight markers, and for somatic or germline
 CC gene therapy. The present sequence represents a dichp protein of the
 CC invention. Note: The sequence data for this patent is not represented in
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
 XX
 SQ Sequence 1260 AA;
 Query Match 12.4%; Score 73.5; DB 8; Length 1260;
 Best Local Similarity 24.2%; Pred. No. 26;
 Matches 23; Conservative 21; Mismatches 32; Indels 19; Gaps 5;
 QY 18 HPLAQKRGARILVFTDDPRRSVLIVPGCHLDSMRREKNAYYFQDGNALVGMVSGGTV-- 75
 DB 589 HRLSTVRNADVIAGFED---GVIVEQSGHSELMKKE--GVYFK---LVNMQTSGSQIQS 639
 QY 76 -EYDADDRTYVVQLTDG-----RHTTSEFSFHS 102
 DB 640 EEFELNDEKAATRMAPNGWKSRLFRHSTQKLNKS 674
 RESULT 10
 ID ADP56345 standard; protein; 1286 AA.
 XX ADP56345;
 AC ADP56345;
 XX 18-NOV-2004 (first entry)
 DT Human PRO protein sequence SEQ ID NO:2321.
 DE human; PRO; immune related disease; inflammatory immune response;
 KW immune response stimulation; anti-allergic; antianaemic; antiarthritic;
 KW antiasthmatic; antidiabetic; antiinflammatory; antipsoriatic;
 KW antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;
 KW haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
 KW nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
 KW virucide; gene therapy.
 XX
 OS Homo sapiens.
 XX WO2004039956-A2.
 PN 13-MAY-2004.
 PD 28-OCT-2003; 2003WO-US034381.
 XX 29-OCT-2002; 2002US-0422472P.
 PR (GETH) GENENTECH INC.
 XX Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;
 PI Wood WI, Wu TD;
 XX WPI; 2004-376182/35.
 DR N-PSDB; ADP56344.
 XX New PRO polynucleotides and polypeptides, useful in diagnosing

PT and treating an immune related disease, e.g. systemic lupus
 PS erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
 PT stimulating an immune response.
 XX
 PS Claim 1; SEQ ID NO 2321; 3009pp; English.
 XX
 CC The present invention describes an isolated PRO nucleic acid (1). Also
 CC described: (1) a vector comprising (1); (2) a host cell comprising the
 CC vector of (1); (3) a process for producing a PRO polypeptides; (4) an
 CC isolated PRO polypeptide; (5) a chimeric molecule comprising the
 CC polypeptide of (4) fused to a heterologous amino acid sequence; (6) an
 CC antibody which specifically binds to a polypeptide of (4); (7) a
 CC composition of matter comprising a polypeptide of (4); an agonist or
 CC antagonist of the polypeptide or an antibody that binds to the
 CC polypeptide in combination with a carrier; (8) an article of manufacture
 CC comprising a container, a label on the container and a composition of
 CC matter of (7); (9) a method of treating an immune related disease in a
 CC mammal; (10) a method for determining the presence of a PRO polypeptide
 CC in a sample suspected of having the polypeptide; (11) a method of
 CC diagnosing an immune related disease or an inflammatory immune response
 CC in mammal; (12) a method of identifying a compound that inhibits or
 CC mimics the activity of or expression of a gene encoding a PRO polypeptide
 CC ; and (13) a method of stimulating the immune response in a mammal. The
 CC PRO sequences have anti-allergic, antianaemic, antiarthritic,
 CC antiasthmatic, antidiabetic, antiinflammatory, antipsoriatic,
 CC antirheumatic, antithyroid, CNS, dermatological, gastrointestinal,
 CC haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,
 CC nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and
 CC virucide activities, and can be used in gene therapy. The nucleic acid
 CC (1) and the encoded polypeptides, compositions, kits and methods are
 CC useful in diagnosing and treating an immune related disease and in
 CC stimulating an immune response. The present sequence represents a human
 CC PRO protein from the present invention.
 XX
 SQ Sequence 1286 AA;
 Query Match 12.4%; Score 73.5; DB 8; Length 1286;
 Best Local Similarity 24.2%; Pred. No. 26;
 Matches 23; Conservative 21; Mismatches 32; Indels 19; Gaps 5;
 QY 18 HPLAQKRGARILVFTDDPRRSVLIVPGCHLDSMRREKNAYYFQDGNALVGMVSGGTV-- 75
 DB 589 HRLSTVRNADVIAGFED---GVIVEQSGHSELMKKE--GVYFK---LVNMQTSGSQIQS 639
 QY 76 -EYDADDRTYVVQLTDG-----RHTTSEFSFHS 102
 DB 640 EEFELNDEKAATRMAPNGWKSRLFRHSTQKLNKS 674
 RESULT 11
 ID ABM83182 standard; protein; 1294 AA.
 XX ABM83182;
 AC ABM83182;
 XX 18-NOV-2004 (first entry)
 DT Human diagnostic and therapeutic pprotein SEQ ID NO:3431.
 DE gene therapy; human diagnostic and therapeutic polynucleotide; dichp.
 XX Homo sapiens.
 XX WO2004023973-A2.
 PN 25-MAR-2004.
 PD 12-SEP-2003; 2003WO-US028227.
 XX 12-SEP-2002; 2002US-0410259P.
 PR 12-SEP-2002; 2002US-0410260P.
 XX (INCY-) INCYTE CORP.
 PA

XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX WPI: 2004-329368/30.
DR N-PSDB; ACN41834.
XX
PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
PS Claim 27; Page: 190pp; English.
XX
CC The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ Sequence 1294 AA;
Query Match 12.4%; Score 73.5; DB 8; Length 1294;
Best Local Similarity 24.2%; Pred. No. 26;
Matches 23; Conservative 21; Mismatches 32; Indels 19; Gaps 5;
QY 18 HPLAQKRGARILVFTDDPRRSVLIVPGCHLDSMRREKNAYVFDGNALVGMVVGTV-- 75
Db 589 HRLSTVRNADVIAGFED---GVIVEQGHSHSLMKKE--GVYFK---LVNMQTSGSIQIS 639
QY 76 -EYDADRTYVVQLTDG-----RHTTSSFEHS 102
Db 640 EEFELNDEKAATRMAPNGWKSLFRHSTQKNLKNS 674
RESULT 12
ABM83183
ID ABM83183 standard; protein; 1295 AA.
XX
AC ABM83183;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human diagnostic and therapeutic pprotein SEQ ID NO:3432.
XX
KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX Homo sapiens.
XX WO2004023973-A2.
PN
XX 25-MAR-2004.
XX
PD 12-SEP-2003; 2003WO-US028227.
PF
XX 12-SEP-2002; 2002US-0410259P.
PR

PR 12-SEP-2002; 2002US-0410260P.
XX (INCY-) INCYTE CORP.
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX WPI: 2004-329368/30.
DR N-PSDB; ACN41835.
XX
PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
PS Claim 27; Page: 190pp; English.
XX
CC The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ Sequence 1295 AA;
Query Match 12.4%; Score 73.5; DB 8; Length 1295;
Best Local Similarity 24.2%; Pred. No. 26;
Matches 23; Conservative 21; Mismatches 32; Indels 19; Gaps 5;
QY 18 HPLAQKRGARILVFTDDPRRSVLIVPGCHLDSMRREKNAYVFDGNALVGMVVGTV-- 75
Db 589 HRLSTVRNADVIAGFED---GVIVEQGHSHSLMKKE--GVYFK---LVNMQTSGSIQIS 639
QY 76 -EYDADRTYVVQLTDG-----RHTTSSFEHS 102
Db 640 EEFELNDEKAATRMAPNGWKSLFRHSTQKNLKNS 674
RESULT 13
AAU80055
ID AAU80055 standard; protein; 1527 AA.
XX
AC AAU80055;
XX
DT 30-JUL-2002 (first entry)
XX
DE Leuconostoc mesenteroides dextranase.
XX
KW Dextranase; yoghurt; curd; cheese; fermented milk; infant formulae;
KW pet food; vitamin; oral vaccine; enzyme.
XX
OS Leuconostoc mesenteroides.
XX
PN EP1201131-A1.
XX
PD 02-MAY-2002.

XX		23-OCT-2000; 2000EP-00123012.	
XX	PF		
XX	PR	23-OCT-2000; 2000EP-00123012.	
XX		(NEST) SOC PROD NESTLE SA.	
XX	PA		
XX	PI	Bauché A, De Maleprade D, Duboc P, Neubauer H, Zink R;	
XX	PP		
XX	DR	WPI; 2002-373873/41.	
XX	DR	N-PSDB; ARK50932.	
XX	DR		
XX	PT	Preparing fermented food products, e.g. yoghurt, using microorganisms that have dextran sucrose bound to the surface, provide attractive and uniform texture.	
XX	PT		
XX	PS	Disclosures; Page 14-20; 38pp; English.	
XX	PS		
XX	CC	The invention relates to preparation of a fermented food product that includes treating the microorganism (A) to be used for fermentation with a dextranucrase so that this binds to the cell walls of (A). The method is used to produce yoghurt, curd, cheese or other fermented milk products, ice cream, fermented cereal products, water-based jellies, infant formulae and pet foods. Dextranucrase can also be bound to e.g. vitamins or oral vaccines for delivering these to foods, using (A) as carrier. The method produces foods with attractive and uniform texture, since dextranucrase is evenly distributed and forms a thickening agent in situ, with the amount formed being controlled by the amount of dextranucrase bound. Milk does not have to be coagulated in order to achieve a good texture, so the amount of acidic by-products formed is reduced. Dextranucrase binds to many different types of cells over wide ranges of temperature and pH and the treated bacteria can be stored in liquid or powdered form. The present sequence represents the amino acid sequence of dextranucrase	
XX	CC		
XX	SQ	Sequence 1527 AA;	
	Query Match	12.2%; Score 72.5; DB 5; Length 1527;	
	Best Local Similarity	29.2%; Pred. No. 44;	
	Matches	21; Conservative 9; Mismatches 23; Indels 19; Gaps 3;	
Qy	43	PGCH-----LSMRREKNAYFPDGN--ALVGMVVSGGTVEY-----DADDRT 83	
Db	1305	PGCGNLAIRYVQNKSQWFYFDGNGHVAVTGQTINGKKQFYNDGHOSKGFEINADGDT 1364	
Qy	84	VYVOLTDGRHTT 95	
Db	1365	FYTSATDGLRVT 1376	
	RESULT 14		
ID	ADA20755		
XX	ID	ADA20755 standard; protein; 809 AA.	
XX	XX	ADA20755;	
XX	AC		
XX	AC		
DT	20-NOV-2003	(first entry)	
XX	DE	Soybean phospholipase D alpha #2.	
XX	XX		
KW	Enzyme; plant; respiratory burst oxidase; Rboh; tRNA-mnm-s-U-MT;		
KW	chromomethylase; cytosine 5-methyltransferase; phospholipase D;		
KW	transcription factor IIF; asparaginyl tRNA transferase;		
KW	glutaminyl tRNA transferase; EDS; adaptin AP50; adaptin alpha;		
KW	adaptin beta; stress resistance; quality grain improvement; starch;		
KW	herbicide.		
XX	OS	Glycine max.	
XX	OS		
XX	PN	US2003003471-A1.	
XX	PD	02-JAN-2003.	
XX	PD		

Search completed: October 15, 2005, 04:26:19
Job time : 31.0403 secs

Db 236 EISLVRDSRRPKGGQDTGELLKKANEGVKVLMVWDRTSVGLLKKDGLMATHDEET 295
QY 56 AYYFQ-----DGNALV-----GMVWGGTVEYDADRTYVW-- 86
Db 296 AQFFECTEVHCVLCPRNPDDGGSSIVQDLQISTMFTHHQKIVVVDGAMPGEGRRRIVSF 355
QY 87 ----QLTDGRHTT 95
Db 356 VGGIDLCGRYDT 368

RESULT 15
ABB69365
ID ABB69365 standard; protein; 1125 AA.
XX
AC ABB69365;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 34887.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
WPI: 2001-656860/75.
DR N-PSDB; ABL13468.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 34887; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1125 AA;

Query Match 11.9%; Score 71; DB 4; Length 1125;
Best Local Similarity 23.8%; Pred. No. 45;
Matches 25; Conservative 17; Mismatches 37; Indels 26; Gaps 4;
QY 3 EHAIKHMRQRKTKPLAQKRGARILVFTDDPRRSVLIVPGCHLDSMRREKNAYYFDQG 62
Db 599 EHQHQQQQHHHPSHQLG-----LVDECGADEIGAETAGYMGDE 639
QY 63 NALVGMVSGTVEYDAD-DRTYVVLQTDGRHTTSSFEHSSPSR 106
Db 640 DAVCGLLHNG-----YDPDPDSHDHIDH--DPFEGTENNEHHSNGR 678

THIS PAGE BLANK (00070)


```
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratPatSeqFormatter Version 0.9
; SEQ ID NO 976
; LENGTH: 1279
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P21439
US-09-538-092-976

Query Match          12.4%; Score 73.5; DB 4; Length 1279;
Best Local Similarity 24.2%; Pred. No. 5.9;
Matches 23; Conservative 21; Mismatches 32; Indels 19; Gaps 5;

QY 18 HPLAQKRGARILVFTDDPRRSVLIIVPGCHLDSMRREKNAYYFQDGNALVGMVVSQGTVEY 75
Db 589 HRLSTVRNADVIAGFED--GVIVSQGSHSELMKKE--GVYFK----LVNMQTSGSQIQS 639

QY 76 -EYDADRTYVVQLTDG-----RHTTSSFEHS 102
Db 640 BEFELNDEKAATRMAPNGWKSLRFRHSTQKNLKN 674

RESULT 3
US-09-252-991A-17139
; Sequence 17139, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17139
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17139

Query Match          11.8%; Score 70; DB 4; Length 597;
Best Local Similarity 24.6%; Pred. No. 5.7;
Matches 31; Conservative 22; Mismatches 39; Indels 34; Gaps 6;

QY 3 EHAIKRHM-RQRKPTKHPAQKRGARILVFTDDPRRSVLIIVPGCHLDSMRREKNAYYFQD 61
Db 417 BHGFOAHMGSMRSPSR-----GRIHLKSRDPRQH---PSILFNMTMSHEQDWQEFRD 464

QY 62 GNALVGMVVS-----CGTVEYDADRTYVVQLTDGRTHTTSSFEHSPPSR 107
Db 465 GIRLTREIMNOPALDPYGRSLSPGVQSDAELDEFI-----RNHAETAF-HPSCSCK 517

QY 108 PQSDDL 113
Db 518 MGSDDM 523

RESULT 4
US-09-902-540-13844
; Sequence 13844, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
```

```
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13844
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13844

Query Match          11.3%; Score 67.5; DB 4; Length 168;
Best Local Similarity 28.8%; Pred. No. 2;
Matches 32; Conservative 8; Mismatches 24; Indels 47; Gaps 6;

QY 20 LAQKRGAR--ILVFTDDPRRSVLIIVPGCHLDSMRREKNAYYFQDGNALVGMVVSQGTVEY 77
Db 65 LLOPEGARGEVLLAT---YGSVESEPGAHULTLRGKRLHA-----VVVEY 104

QY 78 DADDR-----TYVVO-----LTDG-----RHTTSSFEHSS 103
Db 105 DARDRLAMAGATYVAQAHGAFAEARTVSLLDGPAQCALDDKHSLLELFWEHET 155

RESULT 5
US-09-673-899-2
; Sequence 2, Application US/09673899
; Patent No. 6613335
; GENERAL INFORMATION:
; APPLICANT: Ruelle, Jean Louis
; TITLE OF INVENTION: Polypeptides from Moraxella Branhamella
; TITLE OF INVENTION: Catarrhalis
; FILE REFERENCE: BM45304
; CURRENT APPLICATION NUMBER: US/09/673,899
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: PCT/EP99/02764
; PRIOR FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacteria
US-09-673-899-2

Query Match          11.3%; Score 67.5; DB 4; Length 433;
Best Local Similarity 25.0%; Pred. No. 7.5;
Matches 31; Conservative 13; Mismatches 33; Indels 47; Gaps 5;

QY 25 GARILVFTDDP-----RRSVLI-----VPGCHLDSMRREKNAY 57
Db 307 GIEMVRMTDDPTNIESRSNVIIRRVWQNSPAEHAGLKSQDKIVRIDGVHITS----- 359

QY 58 YFQDGNALVGMV-----SGCTVEYDADRTYVVQLTDGRTHTTSSFE---HS:PSRSPQ 109
Db 360 -----NELVGVVARKAPDSQLTVEIMRDQREMTVOVLAERPSETLSQPQVNSPSQSRQ 414

QY 110 SDDL 113
Db 415 STQL 418

RESULT 6
US-09-673-899-6
; Sequence 6, Application US/09673899
; Patent No. 6613335
; GENERAL INFORMATION:
; APPLICANT: Ruelle, Jean Louis
; TITLE OF INVENTION: Polypeptides from Moraxella Branhamella
; TITLE OF INVENTION: Catarrhalis
; FILE REFERENCE: BM45304
```

```

, CURRENT APPLICATION NUMBER: US/09/673,899
, CURRENT FILING DATE: 2000-12-19
, PRIOR APPLICATION NUMBER: PCT/EP99/02764
, PRIOR FILING DATE: 1999-04-20
, NUMBER OF SEQ ID NOS: 12
, SEQ ID NO 1
, SEQ ID NO 2
, SEQ ID NO 3
, SEQ ID NO 4
, SEQ ID NO 5
, SEQ ID NO 6
, SEQ ID NO 7
, SEQ ID NO 8
, SEQ ID NO 9
, SEQ ID NO 10
, SEQ ID NO 11
, SEQ ID NO 12
, LENGTH: 433
, TYPE: PRT
, ORGANISM: Bacteria
US-09-673-899-6

```

Query Match	11.3%;	Score 67.5;	DB 4;	Length 433;
Best Local Similarity	25.0%;	Pred.No. 7.5;		
Matches	31;	Conservative 13;	Mismatches 33;	Indels 47; Gaps 5;
Qy	25	GARLLVTDDP-----RRSVLI-----	VPGCHLDSMRKNAY	57
Db	307	GIEMVRWTDPTNTIESRNVILIRVQNSPAELGKSGDKIVRGDGHVHTSI-----	359	
Qy	58	YFDGNALVGWVV-----SGGTVEYDADDRYYVVQLTDRGHTTESFE---HSSPSRSPO	109	
Db	360	-----NELGVGARAPDSQLTVFIMRDQRPMTVQVILAERPSETLSQPVNSPSOSRQ	414	
Qy	110	SDDL	113	
Db	415	STOL	418	

```

RESULT 7
US-09-252-991A-21793
; Sequence 21793, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21793
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21793

```

	Query Match	11.3%	Score 67;	DB 4;	Length 188;
	Best Local Similarity	31.7%	Pred. No. 2.7;		
	Matches	19;	Conservative 16;	Mismatches 17;	Indels 8; Gaps 4;
Qy	1	MGEHAIKRHM--RQKPTKTHPLAQRGARILVFTDPRRSVLI--VPGCHLDSMRERKNAY	57		
Db	133	LGSPAARRRLCGRPRRRPOAQR--RLLL--EPRGSVLRCAPARHAAAVORRRRH	187		

RESULT 8
US-09-252-991A-28712
; Sequence 28712, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18

```

; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28712
; LENGTH: 278
; TYPE: prt
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28712

```

Query Match	11.2%	Score 66.5;	DB 4;	Length 278;
Best Local Similarity	23.9%;	Pred. No. 5.4;		
Matches	34;	Conservative 10;	Mismatches 55;	Indels 43; Gaps 5
Qy	4	HAIKRHMQRKPTKTHPLAQRG-----	-----ARIL	29
Db	117	HLHRRORRLURHAHPDAQQQGAGSQRCQQRERPLPPGRRPRHRRGVQQRGAG		176
Qy	30	VFTDDPRRSVLIVPGCHLDGSMREKNAAYFPQDGNALGMVVSQG--TVEYDADDRTYVVQL		88
Db	177	IQVRDPRRS--FVPGSGLCFPPDRRRGRYARPD--LVAQAQPGCVTAAPRPSRGAVLL		231
Qy	89	TGGRHNTTSSFEHSSPSRSPQS		110
Db	232	FAAHHHVTRA---ASPRRAPSA		250

RESULT 9

US-09-252-991A-19274

; Sequence 19274, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 19274

; LENGTH: 401

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-19274

Query Match	11.2%	Score 66.5;	DB 4;	Length 401;
Best Local Similarity	21.4%;	Pred. No. 8.9;		
Matches	28; Conservative 17;	Mismatches 51;	Indels 35;	Gaps 4;
Qy	3	EHAIKRHM-RQRKTHPLQAQKRCARILVTDQP-----RRSVLI-----	41	
Dd	61	EHAMSTNRPERMTAIEIARPCGPEVLVATSRPLPTPGPREVLIVRAAGVGPDVLQR	120	
Qy	42	-----VPGCHLDSMRRE--KNAYYFDGNALGVWVSGGTVEYDADRTTYVVQ	87	
Dd	121	KGYVDPPPGASDIPGLEIAGMVAVGSVERFVAGEAVMALIPGGYAQFAVADERTTLH	180	
Qy	88	LTDGRHTTSS	98	
Dd	181	LPDGLGMEEA	191	

RESULT 10
US-09-270-767-62222
; Sequence 62222, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094

```
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62222
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-62222

Query Match      11.1%; Score 66; DB 4; Length 120;
Best Local Similarity 24.7%; Pred. No. 1.9;
Matches 21; Conservative 14; Mismatches 36; Indels 14; Gaps 2;

QY 3 EHAIKRHRQRKPTKHPKLAQKRGARILVFTDDPRSRLVPGCHLDSMRREKNAYYFQDG 62
DB 25 QHVTGRGHILARQSGKDELVEERGAEVL-----RGKLFAGHDEQGNR-----VQEG 70

QY 63 NALVGMVSGGTVEYDADDRTYVVQ 87
DB 71 ILICAVVFAGGQNGQDAQHAILQ 95

RESULT 11
US-09-270-767-46620
; Sequence 46620, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46620
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-46620

Query Match      11.1%; Score 66; DB 4; Length 157;
Best Local Similarity 24.7%; Pred. No. 2.8;
Matches 21; Conservative 14; Mismatches 36; Indels 14; Gaps 2;

QY 3 EHAIKRHRQRKPTKHPKLAQKRGARILVFTDDPRSRLVPGCHLDSMRREKNAYYFQDG 62
DB 62 QHVTGRGHILARQSGKDELVEERGAEVL-----RGKLFAGHDEQGNR-----VQEG 107

QY 63 NALVGMVSGGTVEYDADDRTYVVQ 87
DB 108 ILICAVVFAGGQNGQDAQHAILQ 132

RESULT 12
US-09-252-991A-27419
; Sequence 27419, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27419
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27419

Query Match      11.1%; Score 66; DB 4; Length 247;
Best Local Similarity 31.5%; Pred. No. 5.3;
Matches 23; Conservative 13; Mismatches 19; Indels 18; Gaps 4;

QY 8 RHMQRKPTKHPKLAQKRGARILVFTDDPRSRLVPGCHLDSMRREKNAYYF-----59
DB 77 RHRQQQP-PHPLNIRGGRRAVRTFNRRPARAGLFVCG-----VRHQHSQPFGARGL 130

QY 60 ----QDGNALVGM 68
DB 131 SLPEREGRAMIGM 143

RESULT 13
US-09-562-737-41
; Sequence 41, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; CURRENT FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-09-562-737-41

Query Match      11.1%; Score 66; DB 4; Length 1024;
Best Local Similarity 25.0%; Pred. No. 38;
Matches 21; Conservative 15; Mismatches 30; Indels 18; Gaps 3;

QY 8 RHMQRKPTKHPKLAQKRGARILVFTDDPRSRLVPGCHLDSMRREKNAYYFQDGNALVG 67
DB 30 RIIRERIP-EAPAGPPSPDFGLFSDDDPKKIWLKALD-----YY-----71

QY 68 MVVSGGTVEYDADDRTYVVQ 91
DB 72 MLRNGDTMEYRKQRPLKIRMLDG 95

RESULT 14
US-08-635-121-2
; Sequence 2, Application US/08635121
; Patent No. 5910442
; GENERAL INFORMATION:
; APPLICANT: Gelman, Irwin H.
; TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,121
; TYPE: PRT
```


03-08-518-2-A-2

	Matches	29; Conservative	9; Mismatches	38; Indels	24; Gaps	3;
QY	5	ATKRHWQRQKPTKHP	LAQKKGARILVFTDD	PRRSVLIVGCGHLD	SMRREKNAYV	----- 58
Db	359	SPKRLVTPRKESK	SLKEEAGRTLV	-----	VGAGCPLSRNVEKNL	GFLPLRNS 408
QY	59	-----	FQGNALVGMVSGGT	VEYDAD	-DRTYVVQLTD	90
Db	409	PDGGRKGWGRFOA	TVTDSGPGVETNE	DEDDVDPAVVP	LSR	448

```

5  A TKHRMRQRKPTKHPPLAKRGNARILVTFDDPSRSLVLPQCHLDSMRREKNAYV----- 58
359 SFKRLVTPRKKSLSKLEEKAGRTL-----VQAGCPFSRNEVEXKHLGFPURNSS 408
59 -----FDGNALVGNVWSGGTVEYDAD--DRTYVVVLTD 90
409 PDGGRKGNGROEATVEDSGPVEINEDEPDPVPAWPISE 448

```

ATTORNEY/AGENT INFORMATION:
NAME: Clark, Richard S
REGISTRATION NUMBER: 26,154
REFERENCE/DOCKET NUMBER: A30558 - 165/34008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2558
TELEFAX: 212-765-2519

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:14:10 ; Search time 27.0227 Seconds
(without alignments)
1742.576 Million cell updates/sec

Title: US-10-617-038-5
Perfect score: 595
Sequence: 1 MGEHAIKRHMQRKPTKHPK.....TTSSFEHSSPSRSPQSDLL 113

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/FCR_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	595	100.0	113	15	US-10-617-038-5
2	78.5	13.2	188	20	Sequence 5, Appli
3	73.5	12.4	241	15	Sequence 26046, A
4	73.5	12.4	664	14	Sequence 224808, A
5	73.5	12.4	1066	9	Sequence 7195, Ap
6	73.5	12.4	1275	15	Sequence 5, Appli
7	73.5	12.4	1279	15	Sequence 49, Appli
8	72	12.1	177	15	Sequence 81, Appli
9	71	11.9	254	16	Sequence 59508, A
10	71	11.9	277	16	Sequence 161172, A
11	71	11.9	345	16	Sequence 203069, A
					Sequence 154543, A

12	71	11.9	629	15	US-10-425-114-43053	Sequence 43053, A
13	71	11.9	809	14	US-10-078-770-124	Sequence 124, App
14	71	11.9	809	15	US-10-424-599-206877	Sequence 206877, A
15	71	11.9	1125	20	US-11-097-143-34887	Sequence 34887, A
16	70.5	11.8	654	16	US-10-437-963-104798	Sequence 104798, A
17	70	11.8	561	9	US-09-815-242-12101	Sequence 12101, A
18	70	11.8	561	15	US-10-282-122A-66833	Sequence 66833, A
19	70	11.8	751	16	US-10-476-924-22	Sequence 22, Appli
20	69.5	11.7	275	17	US-10-732-923-8134	Sequence 8134, Ap
21	69.5	11.7	459	15	US-10-425-114-64683	Sequence 64683, A
22	69.5	11.7	459	16	US-10-425-115-358064	Sequence 358064, A
23	69.5	11.7	603	17	US-10-732-923-8095	Sequence 8095, Ap
24	69.5	11.7	808	17	US-10-732-923-8095	Sequence 8095, Ap
25	69.5	11.7	821	17	US-10-732-923-8091	Sequence 8091, Ap
26	69	11.6	216	15	US-10-425-114-50321	Sequence 50321, A
27	68.5	11.5	286	9	US-09-738-626-4787	Sequence 4787, Ap
28	68.5	11.5	506	15	US-10-282-122A-61901	Sequence 61901, A
29	68.5	11.5	2920	15	US-10-038-854-71	Sequence 71, Appli
30	68	11.4	736	16	US-10-425-115-267615	Sequence 267615, A
31	68	11.4	753	15	US-10-425-114-58733	Sequence 58733, A
32	68	11.4	808	15	US-10-424-599-171300	Sequence 171300, A
33	68	11.4	1499	16	US-10-437-963-126416	Sequence 126416, A
34	67.5	11.3	433	14	US-10-393-545-2	Sequence 2, Appli
35	67.5	11.3	433	14	US-10-393-545-6	Sequence 6, Appli
36	67	11.3	80	16	US-10-437-963-141579	Sequence 141579, A
37	67	11.3	126	15	US-10-424-599-157549	Sequence 157549, A
38	67	11.3	325	15	US-10-425-114-37607	Sequence 37607, A
39	67	11.3	325	16	US-10-437-963-160059	Sequence 160059, A
40	67	11.3	377	16	US-10-425-115-206647	Sequence 206647, A
41	67	11.3	391	16	US-10-739-930-8398	Sequence 8398, Ap
42	67	11.3	562	16	US-10-425-115-295640	Sequence 295640, A
43	67	11.3	597	16	US-10-425-115-297734	Sequence 297734, A
44	67	11.3	813	15	US-10-425-114-70103	Sequence 70103, A
45	67	11.3	817	15	US-10-425-114-47265	Sequence 47265, A

ALIGNMENTS

RESULT 1
US-10-617-038-5
; Sequence 5, Application US/10617038
; Publication NO. US20040057963A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Peter
; APPLICANT: Rosenkands, Ida
; APPLICANT: Stryhn, Anette
; TITLE OF INVENTION: Therapeutic TB Vaccine
; FILE REFERENCE: SS15AUSA
; CURRENT APPLICATION NUMBER: US/10/617,038
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: DK PA 2002 01098
; PRIOR FILING DATE: 2002-07-13
; PRIOR APPLICATION NUMBER: US 60/401,725
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-617-038-5

Query Match	100.0%	Score 595;	DB 15;	Length 113;
Best Local Similarity	100.0%	Pred. No. 2.2e-64;		
Matches 113;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MGEHAIKRHMQRKPTKHPKLAQKRGARILVFDDPRRSVLIVPGCHLDSMRREKNAYYFQ	60	
DB	1	MGEHAIKRHMQRKPTKHPKLAQKRGARILVFDDPRRSVLIVPGCHLDSMRREKNAYYFQ	60	
QY	61	DGNALGVVSGTVEYDADRTYVQLTDGRTTSSFEHSSPSRSPQSDLL	113	

Db 61 DGNALVGMVSGTVEYDADRTYVVLTDGRHTTSSFEHSSPSRSPQSDDL 113

```

RESULT 2
; Sequence 26046, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26046
; LENGTH: 188
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-10-097-143-26046

Query Match 13.2%; Score 78.5; DB 20; Length 188;
Best Local Similarity 30.4%; Pred. No. 0.56;
Matches 34; Conservative 14; Mismatches 45; Indels 19; Gaps 6;

Qy 1 MGEHAIKHMRQKTKPLAQKRGARILVTDPRRSVLVPGCHLDSMR---EKVAY 57
Db 43 MFAHKPSTHKRQYLTPNPDPSRSA-TDFDFDRPSVLEIP-----DRSYRLVNLRLAY 97
Qy 58 YFQDGNALVGMVSGTVEYDADD---RTYVVLQ-----TDGRHTTSSFE 100
Db 98 GLLDSGASISCI--GGQLEKAATEPENFKSLTIELLTKTKFGRSSSDSFE 147

RESULT 3
US-10-424-599-224808
; Sequence 224808, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 224808
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
```

```

; OTHER INFORMATION: Clone ID: PAT_MRT3847_45032C.1.pcp
US-10-424-599-224808

Query Match 12.4%; Score 73.5; DB 15; Length 241;
Best Local Similarity 30.6%; Pred. No. 3.2;
Matches 26; Conservative 16; Mismatches 32; Indels 11; Gaps 4;

Qy 20 LAQKRGARILVTDPR-----RSVLVPGCHLDSMRREKNA--YFQDGNALVGMVSG 72
Db 103 ISMRGASIGWHSDNRPYLKQRHFSAV--CYLNTYKGDFSGGLFHPQDGEPKSIVPKAG 160
Qy 73 GTVEYDADRTY--VWOLTDGRHTT 95
Db 161 DVVMYTADRNHISVDEITEGERLT 185

RESULT 4
US-10-032-585-7195
; Sequence 7195, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7195
; LENGTH: 664
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7195

Query Match 12.4%; Score 73.5; DB 14; Length 664;
Best Local Similarity 24.1%; Pred. No. 12;
Matches 27; Conservative 21; Mismatches 49; Indels 15; Gaps 4;

Qy 12 QRKPTKPLAQKRGARILVTDPRRSVLVPGCHLDSMRREKNAYFQDG---NALVG 67
Db 74 QTKLAKNSTPVSCKTRDTIGSTSNRKQFRTVDNHHLLTPQLKKRK---FEDNTPEKTPAG 130
Qy 68 MVVSGGTVEYDADRTYVVLTDGRHTTSSFEHSSPSRS-----PQSDDL 113
Db 131 IPVSSPTDYETANNFTVGLTTNGTPTTTTA--KTPFSNLSLETLPNQIDEI 180

RESULT 5
US-09-817-762-5
; Sequence 5, Application US/09817762
; Publication No. US20020068344A1
; GENERAL INFORMATION:
; APPLICANT: Spalding, Edgar P.
; APPLICANT: Noh, Bosh
; TITLE OF INVENTION: MDR-Like ABC Transporter Gene From
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 13238-00061
; CURRENT APPLICATION NUMBER: US/09/817,762
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: PCT/US99/22363
; PRIOR FILING DATE: 1999-09-24
; PRIOR APPLICATION NUMBER: US 60/101,814
; PRIOR FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1066
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
```

```
; DATABASE ACCESSION NUMBER: Genbank P21439
; DATABASE ENTRY DATE: 1998-07-15
US-09-817-762-5

Query Match          12.4%; Score 73.5; DB 9; Length 1066;
Best Local Similarity 24.2%; Pred. No. 23; Mismatches 32; Indels 19; Gaps 5;
Matches 23; Conservative 21;

QY 18 HPLAQRGARILVFTDDPRRSVLIVPGCHLDSMRREKNAYFFQDGNALYGMVVSQGTG-- 75
Db 376 HRLSTVRNADVIAGPED---GVIVEGSHSELMKKE--GVYFK----LVNMQTSSGSIQS 426
QY 76 -EYDADRTYVVQLTDG-----RHTTESSFEHS 102
Db 427 BEFELNDEKAATRMAPNGWKSRLFRHSTQKNLKN 461

RESULT 6
US-10-363-112-49
; Sequence 49, Application US/10363112
; Publication No. US20040091964A1
; GENERAL INFORMATION:
; APPLICANT: THE AUSTRALIAN NATIONAL UNIVERSITY
; APPLICANT: BOARD, PHILLIP
; APPLICANT: HARRIS, MATTHEW
; TITLE OF INVENTION: MODIFIED PROTEINS, ISOLATED NOVEL PEPTIDES, AND USES THEREOF
; FILE REFERENCE: 007643-0302189
; CURRENT APPLICATION NUMBER: US/10/363,112
; CURRENT FILING DATE: 2003-11-03
; PRIOR APPLICATION NUMBER: PCT/AU01/01093
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/229,663
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
; LENGTH: 1275
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-363-112-49

Query Match          12.4%; Score 73.5; DB 15; Length 1275;
Best Local Similarity 24.2%; Pred. No. 29; Mismatches 32; Indels 19; Gaps 5;
Matches 23; Conservative 21;

QY 18 HPLAQRGARILVFTDDPRRSVLIVPGCHLDSMRREKNAYFFQDGNALYGMVVSQGTG-- 75
Db 589 HRLSTVRNADVIAGPED---GVIVEGSHSELMKKE--GVYFK----LVNMQTSSGSIQS 639
QY 76 -EYDADRTYVVQLTDG-----RHTTESSFEHS 102
Db 640 BEFELNDEKAATRMAPNGWKSRLFRHSTQKNLKN 674

RESULT 7
US-10-618-281-81
; Sequence 81, Application US/10618281
; Publication No. US20040219609A1
; GENERAL INFORMATION:
; APPLICANT: Day, Anthony G.
; APPLICANT: Estell, David A.
; APPLICANT: Lyons, Eric H.
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: Methods for Modulating Proteins Not
; FILE REFERENCE: GC73-2
; CURRENT APPLICATION NUMBER: US/10/618,281
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: US 60/395,325
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
```

```
; LENGTH: 1279
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-618-281-81

Query Match          12.4%; Score 73.5; DB 16; Length 1279;
Best Local Similarity 24.2%; Pred. No. 29; Mismatches 32; Indels 19; Gaps 5;
Matches 23; Conservative 21;

QY 18 HPLAQRGARILVFTDDPRRSVLIVPGCHLDSMRREKNAYFFQDGNALYGMVVSQGTG-- 75
Db 589 HRLSTVRNADVIAGPED---GVIVEGSHSELMKKE--GVYFK----LVNMQTSSGSIQS 639
QY 76 -EYDADRTYVVQLTDG-----RHTTESSFEHS 102
Db 640 BEFELNDEKAATRMAPNGWKSRLFRHSTQKNLKN 674

RESULT 8
US-10-425-114-59508
; Sequence 59508, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59508
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3355-052-B10_FLI.pep
US-10-425-114-59508

Query Match          12.1%; Score 72; DB 15; Length 177;
Best Local Similarity 22.9%; Pred. No. 3.2; Mismatches 20; Indels 34; Gaps 5;
Matches 27; Conservative 20;

QY 2 GEHAKRHWQRKPTKHPKLAQKRGARILVFTDDPRRSVLIVPGCHLDSMRREKNAYFFQD 61
Db 29 GRLLRLRLRLHRRPPRPVQVKHARVPLVDPDPPDARLINLP--HL----- 72
QY 62 GNALYGMV-----SGGTG-----EYDADRTYVVQLTD-----GRHTTESSFEHS 102
Db 73 -RAFLGLICRWAELESAGTTAFTPLHQRETDLSGAVQLPDPSSAAGRGVVCFCFAYA 129

RESULT 9
US-10-437-963-161172
; Sequence 161172, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
```


RESULT 13

RESULT 15

RESULT 14

Query Match 11.9%; Score 71; DB 20; Length 1125;
Best Local Similarity 23.8%; Pred. No. 50;
Matches 25; Conservative 17; Mismatches 37; Indels 26; Gaps 4;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:59:30 ; Search time 6.3053 Seconds
(without alignments)
1724.342 Million cell updates/sec

Title: US-10-617-038-5
Perfect score: 595
Sequence: 1 MGEHAIKRHMQRKPTKHPLE.....TTSESPFHSRSPSQSDDL 113

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: piri:.*
2: pirt:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	595	100.0	113	2	C709333
2	74.5	12.5	324	2	S44956
3	73.5	12.4	576	2	H82777
4	73.5	12.4	1279	1	DVHU3
5	73	12.3	383	2	A96704
6	72	12.1	284	2	T47277
7	70	11.8	561	2	B82975
8	69.5	11.7	603	2	JC4666
9	69	11.6	342	2	T22742
10	68.5	11.5	686	2	A59348
11	68	11.4	326	2	E71602
12	67.5	11.3	833	2	A84513
13	67	11.3	353	1	H71340
14	66	11.1	547	2	S19607
15	66	11.1	1174	2	S28976
16	66	11.1	1212	2	C82834
17	66	11.1	1346	2	A57376
18	65.5	11.0	449	2	T49125
19	65.5	11.0	794	2	C87437
20	65.5	11.0	875	1	URXLA2
21	65.5	11.0	1438	2	A48216
22	65.5	11.0	1471	2	B48218
23	65.5	11.0	1578	2	I48216
24	65	10.9	402	1	B24706
25	65	10.9	402	2	B95320
26	64.5	10.8	249	2	G70253
27	64.5	10.8	381	2	S46583
28	64.5	10.8	457	2	G72053
29	64.5	10.8	457	2	B86569

30	64.5	10.8	462	2	B81613
31	64.5	10.8	489	2	E86366
32	64.5	10.8	593	2	AF1011
33	64.5	10.8	801	2	T07617
34	64.5	10.8	890	2	G64740
35	64.5	10.8	1469	2	T19168
36	64.5	10.8	1592	2	A38175
37	64	10.8	312	2	H64162
38	64	10.8	432	2	A28913
39	64	10.8	697	2	G70133
40	64	10.8	1280	1	DVHU1
41	63.5	10.7	227	2	S35564
42	63.5	10.7	231	2	T31231
43	63.5	10.7	323	2	D83608
44	63.5	10.7	357	1	JC1256
45	63.5	10.7	357	1	MNXRS2

ALIGNMENTS

RESULT 1

C70933
hypothetical protein Rv0572c - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: C70933
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: C70933
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-113 <COL>
A;Cross-references: UNIPROT:O53769; GB:AL021942; GB:AL123456; NID:G3242298; PIDN:CAA1744; C;Experimental source: strain H37RV
C;Genetics:
A;Gene: Rv0572c
C;Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0572c

Query Match 100.0%; Score 595; DB 2; Length 113;
Best Local Similarity 100.0%; Pred. No. 2.6e-56;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGEHAIKRHMQRKPTKHPLEAKRGARILVFTDDPRSVLIVPGCHLDSMRREKNAYFQ 60

Db 1 MGEHAIKRHMQRKPTKHPLEAKRGARILVFTDDPRSVLIVPGCHLDSMRREKNAYFQ 60

Qy 61 DGNALVGMVVGCTGVYDADRTYVVLQTDGRTHTTSSPFHSRSPSQSDDL 113

Db 61 DGNALVGMVVGCTGVYDADRTYVVLQTDGRTHTTSSPFHSRSPSQSDDL 113

RESULT 2

S44956
lmbI protein - Streptomyces lincolnensis
C;Species: Streptomyces lincolnensis
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S69818; S44956
R;Peschke, U.; Schmidt, H.; Zhang, H.Z.; Piepersberg, W. Mol. Microbiol. 16, 1137-1156, 1995
A;Title: Molecular characterization of the lincomycin-production gene cluster of Streptomyces lincolnensis
A;Reference number: S69805; MUID:96020646; PMID:8577249
A;Accession: S69818
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-324 <PE2>
A;Cross-references: UNIPROT:Q54362; EMBL:X79146; NID:g499194; PIDN:CAA55755.1; PID:g5816
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1994

C:Genetics:
A:Gene: lmb1
A:Start codon: GTG

Query Match 12.5%; Score 74.5; DB 2; Length 324;
Best Local Similarity 32.0%; Pred. No. 3.1;
Matches 33; Conservative 9; Mismatches 36; Indels 25; Gaps 7;

QY 17 KHPLAQKRGARILVTDPRRSVLVPGC-HLD---SMRREKNAYVFQDGNALVGMVVS 71
DB 122 REELASEAAR-----ALAVPGCDHAEALSVHRRTHLLFADGGDLAKESVQS 170

QY 72 GTVEYDADD-----RTYVOLTQGRHTTESSEFSHSPSRSP 108

DB 171 GQWVRATADPGSPGIRIYPER--TGRHAS-AGWEHIAALRLP 210

RESULT 3

H82777
alkaline phosphatase XF0657 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: H82777
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: H82777
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-576 <SIM>
A:Cross-references: UNIPROT:Q9PFX0; GB:AE003910; GB:AE003849; NID:g9105532; PIDN:APF8346
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Teuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328

C:Contents: annotation

C:Genetics:

A:Gene: XF0657

C:Superfamily: alkaline phosphatase

Query Match 12.4%; Score 73.5; DB 2; Length 576;
Best Local Similarity 28.3%; Pred. No. 7.7;
Matches 26; Conservative 7; Mismatches 56; Indels 3; Gaps 2;

QY 4 HAIKHMRORRKTUP-LAQKRGARILVTDPRRSVLVPGCHLDSMRREKNAYVFQD 62

DB 314 HMQFEHDDRHPNGEFSLAEMRTATQSLSRDPGHGFLMVEGGRIDHAHHAGNAYRALDE 373

QY 63 NALVGMVSGGVVEYDADDRTYVOLTQGRHT 94

DB 374 TIALSDAVRAAVQSPKQ--TLIIVTADHAHT 403

RESULT 4

DVHU3

multidrug resistance protein 3 - human

N:Alternate names: P-glycoprotein MDR3

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004

A:Accession: JS0051; S01346; A42213; I38015

R:van der Bliek, A.M.; Koolman, P.M.; Schneider, C.; Borst, P.

Gene 71, 401-411, 1988
A:Title: Sequence of mdr3 cDNA encoding a human P-glycoprotein.
A:Reference number: JS0051; MUID:89138016; PMID:2906314

A:Accession: JS0051

A:Molecule type: mRNA

A:Residues: 1-1279 <VAL>

A:Cross-references: UNIPROT:P21439; GB:M23234; NID:g187501; PIDN:AAA36207.1; PID:g307181

R:van der Bliek, A.M.; Baas, F.; ten Houte de Lange, T.; Koolman, P.M.; van der Velde-Koe

ENBO J. 6, 3325-3331, 1987

A:Title: The human mdr3 gene encodes a novel P-glycoprotein homologue and gives rise to

A:Reference number: S01346; MUID:88111519; PMID:2892668

A:Accession: S01346

A:Molecule type: mRNA

A:Residues: 856-1093 'FVDFGFGQ' 1094-1279 <VA2>

A:Cross-references: EMBL:X06181; NID:934524; PIDN:CAA29547.1; PID:g34525

R:Lincke, C.R.; Smit, J.J.M.; van der Velde-Koerts, T.; Borst, P.

J. Biol. Chem. 266, 5303-5310, 1991

A:Title: Structure of the human MDR3 gene and physical mapping of the human MDR locus.

A:Reference number: A42213; MUID:91161629; PMID:2002063

A:Accession: A42213

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-1279 <LIN>

R:Smit, J.J.; Mol, C.A.; van Desmter, L.; Wagenaar, E.; Schinkel, A.H.; Borst, P.

Biochim. Biophys. Acta 1261, 44-56, 1995

A:Title: Characterization of the promoter region of the human MDR3 P-glycoprotein gene.

A:Reference number: I38015; MUID:95200972; PMID:7893760

A:Accession: I38015

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-61 'RGSRRVDLQAC' <RES>

A:Cross-references: EMBL:X35284; NID:g1006662; PIDN:CAA84542.1; PID:g1006663

C:Comment: This is an integral membrane protein overproduced in multidrug-resistant cell

structurally and functionally unrelated lipophilic anticancer drugs.

C:Genetics:

A:Gene: GDB:PGY3; MDR3

A:Cross-references: GDB:120713; OMIM:171060

A:Map position: 7q21-7q21

C:Superfamily: multidrug resistance protein; ATP-binding cassette homology

C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; transmembrane pr

F:1-640,653-1279/Region: duplication

F:412-606/Domain: ATP-binding cassette homology <ABC1>

F:429-436/Region: nucleotide-binding motif A (P-loop)

F:553-557/Region: nucleotide-binding motif B

F:638-694/Domain: linker <LINK>

F:1052-1248/Domain: ATP-binding cassette homology <ABC2>

F:1069-1076/Region: nucleotide-binding motif A (P-loop)

F:1195-1199/Region: nucleotide-binding motif B

F:91.97/Binding site: carboxylate (Asn) (covalent) #status predicted

F:435/Binding site: ATP (Lys) #status predicted

F:1075/Binding site: ATP (Lys) #status predicted

Query Match 12.4%; Score 73.5; DB 1; Length 1279;
Best Local Similarity 24.2%; Pred. No. 19;
Matches 23; Conservative 21; Mismatches 32; Indels 19; Gaps 5;

QY 18 HPLAQKRGARILVTDPRRSVLVPGCHLDSMRREKNAYVFQDGNALVGMVSGTV-- 75

DB 589 HRLSTRNADYIAGFED---GVIVEQGSLSHLMKKE--GVYFK---LVNMQTSQSIQS 639

QY 76 -EYDADRTYVVLTDG-----RHTTESSFEHS 102

DB 640 EEFELNDEKAATRNAPNGWKSRLFRHSTQKLNKS 674

hypothetical protein T23K23.7 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

A:Accession: A96704

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Vencer, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A96704
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-383 <STO>
A:Cross-references: UNIPROT:Q9C9X2; GB:AE005173; NID:g6553888; PIDN:AAF16554.1; GSPDB:GN
C:Genetics:
A:Gene: T23K23.7
A:Map position: 1

Query Match 12.3%; Score 73; DB 2; Length 383;
Best Local Similarity 29.5%; Pred. No. 5.5;
Matches 28; Conservative 14; Mismatches 37; Indels 16; Gaps 5;
QY 24 RGAAILVFTDPRSVL-----IVPGCHLDSMRRE--KNAYYPQDGNALVGMVVGCTVEY 77
DB 86 KGASIGWSD--NRSYLKORDFAAVCYLNSYKDFTGGLFRFQSGRPVTVAPSAGDVIMY 144
QY 78 DADDRTY--VVQLTDGRHTT-----ESSFEHSS 103
DB 145 TADDRNIHSVDEVTGDERUTLALWFRSDSHDEDS 179

RESULT 6
T47277
lactoylglutathione lyase (EC 4.4.1.5) [imported] - wheat (fragment)
N:Alternate names: glyoxalase I
C:Species: Triticum aestivum (common wheat)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
R:Johansen, K.S.; Svendsen, I.; Rasmussen, S.K.
A:Description: Purification and cloning of a glyoxalase I from wheat bran.
A:Reference number: Z24451
A:Accession: T47277
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-284 <JOH>
A:Cross-references: UNIPROT:Q9XGE2; EMBL:AJ243528; PIDN:CAB50787.1
A:Experimental source: bran
C:Genetics:
A:Gene: lgul
C:Keywords: carbon-sulfur lyase

Query Match 12.1%; Score 72; DB 2; Length 284;
Best Local Similarity 21.3%; Pred. No. 5;
Matches 29; Conservative 18; Mismatches 37; Indels 52; Gaps 5;
QY 14 KPTKHPLAQKRGARILVFTDDP---RRSVLIVPG-----C 45
DB 115 KVTREPGPVKGGKTVIAFTEDDPGKFEILRGTEPLCQVLMRVGDLDRASFYEKAC 174
QY 46 HLDMSMRREKNAYYF-----QDGNALVGMVVGCTVEYD-----ADRTY- 84
DB 175 GMKLLAKRDNPYKYITAMMGYPEDQNAVLEITYGVTEYDKGNAYQAIGTDDVYK 234
QY 85 ---VVQLTDGRHTTES 97
DB 235 TAEVVKLSGGKVVREA 250

RESULT 7
B82975
choline dehydrogenase PA5372 [imported] - *Pseudomonas aeruginosa* (strain PA01)
C:Species: *Pseudomonas aeruginosa*

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 16-Aug-2004
C:Accession: B82975
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bri-
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho-
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: B82975
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-561 <STO>
A:Cross-references: UNIPROT:Q9HTJ2; GB:AE004949; GB:AE004091; NID:g9951690; PIDN:AAG0875;
A:Experimental source: strain PA01
C:Genetics:
A:Gene: betaA; PA5372
C:Superfamily: Alcohol oxidase

Query Match 11.8%; Score 70; DB 2; Length 561;
Best Local Similarity 24.6%; Pred. No. 18;
Matches 31; Conservative 22; Mismatches 39; Indels 34; Gaps 6;
QY 3 EHAIKRHM-RQKPKHPLAQKRGARILVFTDDPRSVLIVPGCHLDSMRREKNAYYFQD 61
DB 381 EHGFOAHGSMRSPSR-----GRHLKSRDPQH-----PSILFNYSHEQDWQEFRD 428
QY 62 GNALVGMVVS-----GGTVEYDADRTYVVQLTDGRHTTESSEHSPSPRS 107
DB 429 GIRLREIMNQPALDPYRGRELSPGVSVQSDAELDEFI-----RNHAETAF-HPSCSCK 481
QY 108 POSDDL 113
DB 482 MGSDDM 487

RESULT 8
JC4666
lymphocyte specific helicase - mouse
N:Alternate names: helix destabilizing enzyme
C:Species: Mus musculus (house mouse)
C:Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 09-Jul-2004
C:Accession: JC4666
R:Jarvis, C.D.; Geiman, T.; Vila-Storm, M.P.; Osipovich, O.; Akella, U.; Candeaas, S.; Ne-
Gene 169, 203-207, 1996
A:Title: A novel putative helicase produced in early murine lymphocytes.
A:Reference number: JC4666; MUID:96194802; PMID:8647447
A:Molecule type: mRNA
A:Residues: 1-603 <JAR>
A:Cross-references: UNIPROT:Q60948; GB:U25691; NID:g805295; PIDN:AAB08015.1; PID:g805296
A:Experimental source: fetal thymus
C:Comment: This enzyme is involved in DNA strand separation, replication, repair, recombi-
C:Genetics:
A:Gene: lysh
C:Keywords: ATP; thymus
F:8-26/Region: ATP-binding

Query Match 11.7%; Score 69.5; DB 2; Length 603;
Best Local Similarity 23.0%; Pred. No. 22;
Matches 20; Conservative 14; Mismatches 28; Indels 25; Gaps 2;
QY 19 PLAQKRGARILVFTDDPRSVLIVPGCHLDSM-----RREKNAYYF----- 59
DB 374 PELKKRGHKVLVFSQWTSMLDMDYCHLRNFIISRLDGSMSYSEREKNIYSFNTDPDVF 433
QY 60 -----QDGNALVGMVVGCTVEYDAD 80
DB 434 LFLVSTRAGGLGILNTAADTVIYOSD 460

RESULT 9
T22742
hypothetical protein F55G11.5 - *Caenorhabditis elegans*

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T22742
R:Baynes, C.
submitted to the EMBL Data Library, November 1996
A:Reference number: 219608
A:Accession: T22742
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-342 <WIL>
A:Cross-references: UNIPROT:Q9XUE0; EMBL:282272; PIDN:CAB05216.1; GSPDB:GN000022; CESP:FS
A:Experimental source: clone F55G11
C:Genetics:
A:Gene: CESP:F55G11.5
A:Map position: 4
A:Introns: 19/1; 94/2; 131/1; 199/2; 239/1; 287/3; 308/2

Query Match 11.6%; Score 69; DB 2; Length 342;
Best Local Similarity 28.2%; Pred. No. 13;
Matches 24; Conservative 10; Mismatches 35; Indels 16; Gaps 4;

QY 17 KHPLAQKRGARILVFTDPR-----RSVL-----VPGC-----HLDNRREKNAYYFDQG 62
DB 156 QSIYSRGGIHLPPQDPNTYFSLSALVFEFGSGPGCNVGNLYQMYRSKKPYFSSE 215

QY 63 NALV--GMVSGGTVEYDADRTYV 85
DB 216 GSIIVFNLAAGNSDKLLIQDTYV 240

RESULT 10
A59348
formate dehydrogenase (EC 1.2.1.2) alpha chain MJ1351 [similarity] - Methanococcus janna
C:Species: Methanococcus jannaschii
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 06-Oct-2000
A:Accession: A59348; H64468
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; rzon, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: A59348
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-686 <BU12>
A:Cross-references: GB:U67575; GB:L77117; NID:g1591992
A:Note: this ORF is annotated but not translated in GenBank entry U67575, release 117; t
he authors
A:Note: this ORF is annotated in GenBank entry U67575, release 117 as containing a preme
A:Accession: H64468
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 152-686 <BU11>
A:Cross-references: GB:U67575; GB:L77117; NID:g1591992; TIGR:MJ1353; PID:g
A:Note: the original version of GenBank entry U67575, used an incorrect initiation codon
C:Genetics:
A:Gene: MJ1353
C:Superfamily: formate dehydrogenase
C:Keywords: 4Fe-4S; iron-sulfur protein; metalloprotein; molybdenum; molybdopterin; NAD;
F:23-26-30-58/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted
F:60-145/Active site: Lys, His #status predicted
F:144/Binding site: molybdopterin guanine dinucleotide (Cys) (covalent) #status predicted
F:144/Modified site: selenocysteine #status predicted

Query Match 11.5%; Score 68.5; DB 2; Length 686;
Best Local Similarity 22.0%; Pred. No. 32;
Matches 20; Conservative 21; Mismatches 15; Indels 35; Gaps 5;

QY 17 KHPL-----AQKRGARILVFTDPRSVL-----IVPGCHLDSMRREKNAYYFQG 60
DB 184 QHPLIARRIMRAKDGAKIIVI--DPRRTITAKNSDIVLIQIPGTNVALI----- 231

QY 61 DGNALVGMVSGGTVEYDADRTYVVLDTG 91
DB 232 --NAMINVIKENLI-----DKFIKNRTEG 255

RESULT 11
B71602
Probable secreted protein PFB0935W - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
A:Accession: B71602
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
; Perlea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743; PMID:9804551
A:Accession: B71602
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-326 <GAR>
A:Cross-references: GB:AE001428; GB:AE001362; NID:g3845316; PIDN:AACT1976.1; PID:g384531;
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB0935W

Query Match 11.4%; Score 68; DB 2; Length 326;
Best Local Similarity 25.3%; Pred. No. 16;
Matches 21; Conservative 17; Mismatches 17; Indels 28; Gaps 3;

QY 29 LVFTDDPRRSVLIVPGCHLDSMRREKNAYYFDGNALVGMVSGGTVEY----- 77
DB 83 MIGNDELHKNLSILEKILDSLKDK-----LKLPLIKSGTEYLDISKFKKIL 132

QY 78 -DADRTYVVLDTGRRHTTSSSF 99
DB 133 TSDSDKTYLP-----TLESSF 149

RESULT 12
A84513
hypothetical protein At2g14010 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: A84513
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.B.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: A84513
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-833 <STO>
A:Cross-references: UNIPROT:Q9ZPT8; GB:AE002093; NID:g4388830; PIDN:AAID19785.1; GSPDB:GN
C:Genetics:
A:Gene: At2g14010
A:Map position: 2

Query Match 11.3%; Score 67.5; DB 2; Length 833;
Best Local Similarity 26.0%; Pred. No. 51;
Matches 32; Conservative 15; Mismatches 33; Indels 43; Gaps 8;

QY 1 MGEHAIKRH-----MQRQTKHPL-AQKRGARILVFTDPRSVLIVPQCH 46
DB 417 MGVTMKNKSATDILDQVLGEYTLVQRRTLPPEVDTRSGEIAASVEKP-----LSP--Q 469

QY 47 LDSMRREKNAYYFDGNALVGMVSGGTVEYDADRTYVVLDTGRRHTTSSFEHSSPSR 106
DB 470 ADTIAQVK-----GSGIVG-----DSDB---VPG--DGNHDSPPSSNETSI 507

QY 107 SPQ 109

Db 508 APQ 510

RESULT 13

H71340

membrane lipoprotein TmpC precursor - syphilis spirochete

C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)

C:Date: 07-Aug-1998 #sequence_revision 07-Aug-1998 #text_change 09-Jul-2004

C:Accession: H71340; A43595; S29561

R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwirson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDodson, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

Science 281, 375-388, 1998

A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.

A:Reference number: A71250; MUID:98332770; PMID:9665876

A:Accession: H71340

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-353 <COL>

A:Cross-references: UNIPROT:P29724; GB:AE001211; GB:AE000520; NID:g3322582; PIDN:AAC6530

A:Experimental source: strain Nichols

R:Schouls, L.M.; van der Heide, H.G.J.; van Embden, J.D.A.

Infect. Immun. 59, 3536-3546, 1991

A:Title: Characterization of the 35-kilodalton Treponema pallidum subsp. pallidum recombinant antigen p17

A:Reference number: A43595; MUID:91372962; PMID:1894360

A:Accession: A43595

A:Molecule type: DNA

A:Residues: 1-10, 'A', 12-158, 'R', 160-353 <SCH>

A:Cross-references: GB:X57836; NID:g48838; PIDN:CAA40968.1; PID:g581809

A:Note: this protein is shown to incorporate palmitic acid

C:Genetics:

A:Gene: tmpC; TP0319

A:Start codon: GTG

C:Superfamily: ABC transporter yufN

C:Keywords: blocked amino end; lipoprotein; membrane protein; thiolester bond

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-353/Product: membrane lipoprotein tmpC #status predicted <MAT>

F:21/Modified site: fatty acylated amino end (Cys) (in mature form) #status predicted

F:21/Binding site: sn-2,3-diacylglycerol (Cys) (covalent) #status predicted

Query Match 11.3%; Score 67; DB 1; Length 353;

Best Local Similarity 24.74; Pred. No. 22;

Matches 21; Conservative 17; Mismatches 31; Indels 16; Gaps 2;

QY 40 LIVPGCHLDSMRREKNAVYFQDGNALVGMVSGTVEYDADRTYVVLTDG----- 91

Db 16 MLLIGSKSDRPQNGAGEGDFVVGWTDG-----DIDDKSFNQVWEGISRPAQEN 71

QY 92 ----RHTTESFEHSPFSRSPQSD 112

Db 72 NAKCKYVTASTDAEYVPSLSAFAD 96

RESULT 14

S19607

alkaline phosphatase (EC 3.1.3.1) - silkworm

C:Species: Bombyx mori (silkworm)

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Jun-2000

C:Accession: S19607

R:Itch, M.; Takeda, S.; Yamamoto, H.; Izumi, S.; Tomino, S.; Eguchi, M.

Biochim. Biophys. Acta 1129, 135-138, 1991

A:Title: Cloning and sequence analysis of membrane-bound alkaline phosphatase cDNA of the silkworm Bombyx mori

A:Reference number: S19607; MUID:92096456; PMID:1756175

A:Accession: S19607

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-547 <ITO>

A:Cross-references: GB:D90454; NID:g217281; PIDN:BAAL4420.1; PID:g217282

C:Superfamily: alkaline phosphatase

C:Keywords: phosphoric monoester hydrolase

Query Match 11.1%; Score 66; DB 2; Length 547;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:24:25 ; Search time 25.3113 Seconds
(without alignments)
2286.133 Million cell updates/sec

Title: US-10-617-038-5
Perfect score: 595
Sequence: 1 MGEHAIKRHMQRKPTKHL.....TTTSSFFHSPPSRSPQSDLL 113

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 03: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	595	100.0	113	2	O53769	Q53769 mycobacteri
2	589	99.0	113	2	O7D9M1	Q7D9M1 mycobacteri
3	589	99.0	113	2	O7U1R2	Q7U1R2 mycobacteri
4	78.5	13.2	198	2	O7PLD2	Q7PLD2 drosophila
5	78	13.1	678	2	O9AWB8	Q9AWB8 lycopersico
6	78	13.1	807	2	O9FR61	Q9FR61 lycopersico
7	75.5	12.7	1411	2	O73Y53	Q73Y53 mycobacteri
8	74.5	12.5	324	2	O54362	Q54362 streptomyc
9	74.5	12.5	483	2	O9L616	Q9L616 streptomyc
10	73.5	12.4	576	2	O9PFK0	Q9PFK0 xylella fas
11	73.5	12.4	1279	1	MDR3_HUMAN	F21439 homo sapien
12	73	12.3	215	2	O8MHZ3	Q8MHZ3 callithrix
13	73	12.3	383	2	O9C9X2	Q9C9X2 arabidopsis
14	72.5	12.2	516	2	O65KF1	Q65KF1 bacillus li
15	72.5	12.2	546	2	O6CKY2	Q6CKY2 kluveromyc
16	72.5	12.2	552	2	O87BE1	Q87BE1 xylella fas
17	72.5	12.2	715	2	O96CQ9	Q96CQ9 magnaporthe
18	72.5	12.2	1527	2	O8KRE1	Q8KRE1 leuconostoc
19	72	12.1	284	2	O9XGF2	Q9XGF2 triticum ae
20	72	12.1	398	2	O6D882	Q6D882 erwinia car
21	72	12.1	1602	2	O7SEW7	Q7SEW7 neurospora
22	72	12.1	2087	2	O8MXL2	Q8MXL2 leishmania
23	71.5	12.0	730	2	O7RZQ6	Q7RZQ6 neurospora
24	71.5	12.0	1033	2	O6MML9	Q6MML9 bdellovibri
25	71	11.9	290	2	O6ES23	Q6ES23 oryza sativ
26	71	11.9	294	2	O6DEN1	Q6DEN1 brachydanio
27	71	11.9	675	2	O7NWV6	Q7NWV6 chromobacte
28	71	11.9	1125	2	O9W3X0	Q9W3X0 drosophila
29	70.5	11.8	654	2	O8SWH9	Q8SWH9 oryza sativ
30	70.5	11.8	751	2	O9ENL1	Q9ENL1 colorado ti
31	70	11.8	285	2	O04818	O04818 sporobolus

32 70 11.8 561 1 BETA_PSEAE
33 70 11.8 653 1 SCEL_MOUSE
34 69.5 11.7 275 2 O8VDZ1
35 69.5 11.7 568 2 O8P7L4
36 69.5 11.7 808 2 O9CVV7
37 69.5 11.7 821 2 Q60848
38 69.5 11.7 1527 2 Q9ZAR4
39 69.5 11.7 1543 2 Q631L6
40 69 11.6 168 2 Q8L7B6
41 69 11.6 259 2 Q932I0
42 69 11.6 342 2 O9XUE0
43 69 11.6 1505 2 Q8JKF9
44 69 11.6 1832 2 Q6RSW7
45 68.5 11.5 284 2 Q6M608

Q9htj2 pseudomonas
Q9eqg3 mus musculus
Q8vdz1 mus musculus
Q8p7l4 xanthomonas
Q9cvv7 mus musculus
Q60848 mus musculus
Q9zara leuconostoc
Q8l7b6 burkholderi
Q932i0 staphylococ
Q9xue0 caenorhabdi
Q8jxf9 heliothis z
Q6rsw7 banana stre
Q6m608 corynebacte

ALIGNMENTS

RESULT 1

O53769 PRELIMINARY; PRT; 113 AA.
AC O53769
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=Rv0572c;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Roston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
DR EMBL; BX842573; CAA17443.1; -.
DR PIR; C70933; C70933.
DR Tuberculist; Rv0572c; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 113 AA; 12811 MW; 52F7CD86037DF146 CRC64;

Query Match 100.0%; Score 595; DB 2; Length 113;
Best Local Similarity 100.0%; Pred. No. 3.2e-55;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGEHAIKRHMQRKPTKHLAQKRGARILVFTDDPRRSVLIVPGCHLDSMRREKNAYFQ 60
|||||
DB 1 MGEHAIKRHMQRKPTKHLAQKRGARILVFTDDPRRSVLIVPGCHLDSMRREKNAYFQ 60
|||||
QY 61 DGNALGVWVSGTVEYDADRTYVQLTGDGRTTTESSFEHSSPSRSPQSDLL 113
|||||
DB 61 DGNALGVWVSGTVEYDADRTYVQLTGDGRTTTESSFEHSSPSRSPQSDLL 113
|||||

RESULT 2

Q7D9M1 PRELIMINARY; PRT; 113 AA.
AC Q7D9M1
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)

Hypothetical protein.
OrderedLocusNames=MT0599;
OS
Mycobacterium tuberculosis.
NCBI_TaxID=1773;
[1]
RN
SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oghkosh;
RX MEDLINE=22206494; PubMed=12218036;
DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H., Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M., Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and laboratory strains.";
RT J. Bacteriol. 184:5479-5490(2002).
RL EMBL; AE000516; AAK44822.1; -;
DR TIGR; MT0599; -;
KW Hypothetical protein.
SQ SEQUENCE 113 AA; 12777 MW; 08F7C0E60E10914B CRC64;

Query Match 99.0%; Score 589; DB 2; Length 113;
Best Local Similarity 99.1%; Pred. No. 1.4e-54;
Matches 112; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGEHAIKHMRQRKPTKHPLAQKRGARILVFTDDPRRSVLIVPGCHLSMRREKNAYVFQ 60
Db 1 MGEHAIKHMRQRKPTKHPLAQKRGARILVLTDDPRRSVLIVPGCHLSMRREKNAYVFQ 60

Qy 61 DGNALVGMVSGTVEYDADRTYVVQLTDGRHTTSSFEHSPPSRSPQSDDL 113
Db 61 DGNALVGMVSGTVEYDADRTYVVQLTDGRHTTSSFEHSPPSRSPQSDDL 113

RESULT 3
Q7UIR2 PRELIMINARY; PRT; 113 AA.

ID Q7UIR2
AC Q7UIR2;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DN OrderedLocusNames=M50587c;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinomycetaceae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H., Pryor M., Duchoy S., Grondin S., Lacroix C., Monsemp C., Simon S., Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R., Parkhill J., Barrrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RA "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
RW EMBL; BX248335; CAD93449.1; -;
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 113 AA; 12777 MW; 08F7C0E60E10914B CRC64;

Query Match 99.0%; Score 589; DB 2; Length 113;
Best Local Similarity 99.1%; Pred. No. 1.4e-54;
Matches 112; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGEHAIKHMRQRKPTKHPLAQKRGARILVFTDDPRRSVLIVPGCHLSMRREKNAYVFQ 60
Db 1 MGEHAIKHMRQRKPTKHPLAQKRGARILVLTDDPRRSVLIVPGCHLSMRREKNAYVFQ 60

Qy 61 DGNALVGMVSGTVEYDADRTYVVQLTDGRHTTSSFEHSPPSRSPQSDDL 113
Db 61 DGNALVGMVSGTVEYDADRTYVVQLTDGRHTTSSFEHSPPSRSPQSDDL 113

RESULT 4
Q7PLD2 PRELIMINARY; PRT; 198 AA.

ID Q7PLD2
AC Q7PLD2;
DT 01-MAR-2004 (TREMBLrel. 26, Created)
DT 01-WAR-2004 (TREMBLrel. 26, Last sequence update)
DT 01-WAR-2004 (TREMBLrel. 26, Last annotation update)
DE CG17702-PA.
GN Name=CG17702;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=125375568;
RA Celiker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A., Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A., George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R., Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J., Svitek R.S., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C., Weinstein G., Scherer S.B., Myers E.W., Gibbs R.A., Rubin G.M.;
RA "Finishing a whole-genome shotgun: Release 3 of the Drosophila melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:0079-0079(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426071;
RA Hoskins R.A., Smith C.D., Carlson J.W., Carvalho A.B., Halpern A., Kaminker J.S., Kennedy C., Mungall C.J., Sullivan B.A., Sutton G.G., Yasuhara J.C., Wakimoto B.T., Myers E.W., Ceiniker S.E., Rubin G.M., Karpen G.H.;
RA "Heterochromatic sequences in a Drosophila whole-genome shotgun assembly.";
RL Genome Biol. 3:0085-0085(2002).
RN [3]
RP SEQUENCE FROM N.A.
RA Adams M.D., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Li P.W., Henderson S.N., Sutton G.G., Wortman J.H., An H.J., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.H., An H.J., Andrews-Pfannkoch C., Baldwin D., Baller R.M., Basu A., Baxendale J., Beasley E.M., Beeson K.Y., Bhandari D., Bolanos R.A., Busam D.A., Center A., Chandra I., Dahlke C., Davenport L.B., Davies P., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Dunn P., Evangelista C.C., Ferreira S., Flanagan M.J., Foster C., Gabrielian A.E., Garg N.S., Glasser K., Glodek A., Gong F., Gu Z., Guan P., Halpern A.L., Harris M., Heiman T.J., Houck J., Hosdin D., Howland T.J., Wei M.H., Ibegwam C., Jalali M., Kalush F., Ke Z., Ketchum K.A., Kodira C.D., Kraft C., Kravitz S., Lai Z., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McPherson D., Merkulov G., Miller J.R., Milshina N.V., Mobarry C., Moy M., Murphy B., Nelson K.A., Reinstert K., Russek D.R., Pittman G.S., Pan S., Pollard J., Puri V., Reinert K., Remington K., Scheeler F., Shue B.C., Sidon-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Strong R., Sun E., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.Y., Williams S.M., Woodage T., Wu D., Yao Q.A., Ye J., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhu S., Zhu X., Smith H.O., Myers E.W., Venter J.C.;
RA "Drosophila melanogaster Heterochromatic Scaffold.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Li P.W., Henderson S.N., Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.H., An H.J., Andrews-Pfannkoch C., Baldwin D., Baller R.M., Basu A., Baxendale J.,


```

RA Beasley E.M., Beeson K.Y., Bhandari D., Bolanos R.A., Busam D.A.,
RA Center A., Chandra I., Dahlke L., Davenport L.B., Davies P.,
RA Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K.,
RA Doup L.E., Dunn P., Evangelista C.C., Ferriera S., Flanagan M.J.,
RA Foslter C., Gabriellian A.E., Garg N.S., Glaeser K., Glodek A., Gong F.,
RA Gu Z., Guan P., Halpern A.L., Harris M., Heiman T.J., Houck J.,
RA Hostin D., Howland T.J., Wei M.H., Ibegwam C., Jalali M., Kalush F.,
RA Kest Z., Ketchum K.A., Kodira C.D., Kraft C., Kravitz S., Lai Z.,
RA Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X.,
RA Mattei B., McIntosh T.C., McPherson D., Merkulov G., Miller J.R.,
RA Milshina N.V., Mobarry G.S., Pan S., Pollard J., Puri V., Reinert K.,
RA Nussekron D.R., Pittman G.S., Shue B.C., Siden-Kiamos I., Simpson M.,
RA Remington K., Scheeler F., Smith T., Spier E., Strong R., Sun E.,
RA Skupski M.P., Turner E., Wang A.H., Wang X., Wang Z.Y., Williams S.M.,
RA Woodage T., Wu D., Yao Q.A., Ye J.H., Zaveri J.S., Zhan M., Zhang G.,
RA Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhu S., Zhu X.,
RA Smith H.O., Myers E.W., Venter J.C.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Hoskins R.A., Smith C.D., Carlson J.W., Carvalho A.B., Halpern A.,
RA Kaminker J.S., Kennedy C., Lewis S.E., Mungall C.J., Sullivan B.A.,
RA Sutton G.G., Yasuhara J.C., Wakimoto B.T., Myers E.W., Celniker S.E.,
RA Rubin G.M., Karpen G.H.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABU01002199; FAA46285.1; -.
DR InterPro; IPR001969; Pept_AspAS.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN 1.
SQ SEQUENCE 198 AA; 22643 MW; 15CDB1773DF779A3 CRC64;

Query Match 13.2%; Score 78.5; DB 2; Length 198;
Best Local Similarity 30.4%; Pred. No. 4;
Matches 34; Conservative 14; Mismatches 45; Indels 19; Gaps 6;

QY 1 MGEHAIKRHRMRQKPTKHPLAQKRGARILVFTDDPRRSVLIVPGCHLDSMRREKN 57
Db 43 MFAHKSTHKRQVLTTPNDPPSRSA-TFDFDRDPSVLEIP---DRSYRLVNLRLAY 97
QY 58 YFODGNALVGMVSGTVEYDADD-----RTYVVQL-----TDGRHTTESFFE 100
Db 98 GLLDGASISCI--GGQLEKAATEFENFKSLTELKTIKSFGRSSSDSFE 147

RESULT 5
Q9AWB8 PRELIMINARY; PRT; 678 AA.
AC Q9AWB8;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Phospholipase PLDa3 (Fragment).
GN Name=PLDa3;
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21339178; PubMed=11446372;
RA Laxalt A.M., ter Riet B., Verdonk J.C., Parigi L., Tameling W.I.,
RA Vossen J., Haring M., Musgrave A., Munnik T.;
RT "Characterization of five tomato phospholipase D cDNAs: rapid and
RT specific expression of LePLDbeta1 on elicitation with xylanase.";
RL Plant J. 26:237-247(2001).
DR EMBL; AY013254; AAC50297.1; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000008; C2.

Query Match 13.1%; Score 78; DB 2; Length 807;
Best Local Similarity 24.6%; Pred. No. 24;
Matches 33; Conservative 20; Mismatches 41; Indels 40; Gaps 5;

QY 3 EHAIKRHRMRQKRP---TKHPLAQK---GARILVFTDDPRRSVLIVPGCHLDSMRREKN 55
Db 201 EITLIRDMRRPKPGDITLGGELKKKANEGVRVLMVMDRTSVPLVKEDGLMATHDQET 260
QY 56 AYFQ-----DGNALVG-----MVVSGTVEYDADDRTYV-- 85
Db 261 AAYFENSEVHCVLCPRNPDDGRSIIQNIETGTFTHHQKIVVVDGELPNGDTERRRIVSY 320
QY 86 ---VOLTGRHTTE 96
Db 321 IGGIDLCGRYDTQ 334

RESULT 6
Q9FR61 PRELIMINARY; PRT; 807 AA.
AC Q9FR61;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Phospholipase D.
GN Name=TPLD;
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fruit;
RA Whitaker B.D., Smith D.L., Green K.C.;
RT "Cloning, characterization, and functional expression of a
RT phospholipase Da cDNA from tomato fruit.";
RL Physiol. Plantarum 0:0-0(2001).
DR EMBL; AF154425; AAG48162.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004630; F:phospholipase D activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0046470; P:phosphatidylcholine metabolism; IEA.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR001736; PLD.
DR InterPro; IPR011402; PLD_plant.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00614; PLDC; 2.
DR PIRSF; PIRSF036470; PLD_plant; 1.
DR SMART; SM00239; C2; 1.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PS50035; PLD; 2.
SQ SEQUENCE 807 AA; 91984 MW; 65A1C5F36AA56F20 CRC64;

Query Match 13.1%; Score 78; DB 2; Length 807;
Best Local Similarity 24.6%; Pred. No. 24;
Matches 33; Conservative 20; Mismatches 41; Indels 40; Gaps 5;

QY 3 EHAIKRHRMRQKRP---TKHPLAQK---GARILVFTDDPRRSVLIVPGCHLDSMRREKN 55

```

Db 234 EITLIRDMRRPKPGDITLIGELLKKKANEGRVRLVMDRRTSVPLVKEDGLMATHDQET 293

QY 56 AYYFQ-----DGNALVG-----MVVSGTVEYDADRTYV-- 85

Db 284 AAYFENSEVHCVLCRNPDDGRSIIIEIGTMFTTHQKIVVVDGELPNGDTERRRIVSY 353

QY 86 ---VOLTDGRHTTE 96

Db 354 IGGIDLCGRYDTQ 367

RESULT 7

Q73Y53

ID Q73Y53 PRELIMINARY; PRT; 1411 AA.

AC Q73Y53;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE FdhF.

GN Name=fdhF; OrderedLocusNames=MAP2104;

OS Mycobacterium paratuberculosis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI_TaxID=1770;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K10.

RA Li L., Bannantine J., Zhang Q., Anonin A., Alt D., Kapur V.;

RL Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AE017234; AAS04421.1; -

DR HSSP; P00388; 1AM0.

DR GO; GO:0005489; P:electron transporter activity; IEA.

DR GO; GO:0010181; P:FMN binding; IEA.

DR GO; GO:0030151; F:molybdenum ion binding; IEA.

DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR GO; GO:0006118; P:electron transport; IEA.

DR InterPro; IPR009010; Asp decarb fold.

DR InterPro; IPR003097; FAD_binding.

DR InterPro; IPR001094; Flavodoxin-like.

DR InterPro; IPR008254; Flav_nitox_synth.

DR InterPro; IPR001709; FPN_cyt_reductase.

DR InterPro; IPR006656; Molybdopterin.

DR InterPro; IPR006963; Molybdop_Fe4S4.

DR InterPro; IPR006657; Mol_dinuoc_bind.

DR InterPro; IPR001433; Oxid_FAD/NAD(P).

DR InterPro; IPR006655; Prok_Mboxred.

DR Pfam; PF00667; FAD_binding_1; 1.

DR Pfam; PF00258; Flavodoxin_1; 1.

DR Pfam; PF00384; Molybdopterin; 1.

DR Pfam; PF04879; Molybdop_Fe4S4; 1.

DR Pfam; PF01568; Molydop_Binding; 1.

DR Pfam; PF00175; NAD_binding_1; 1.

DR PRINTS; PRO0369; FLAVODOXIN.

DR PRINTS; PRO0371; FPNCR.

DR PROSITE; PS05902; FLAVODOXIN LIKE; 1.

DR PROSITE; PS00551; MOLYBDOPTERIN_PROK_1; UNKNOWN_1.

KW Complete proteome.

SQ SEQUENCE 1411 AA; 153031 MW; ED7212B69C425462 CRC64;

Query Match 12.7%; Score 75.5; DB 2; Length 1411;

Best Local Similarity 24.8%; Pred. No. 84;

Matches 26; Conservative 15; Mismatches 29; Indels 35; Gaps 5;

QY 18 HPL-----AQKRGARILVFTDDRRS-----VLIVPGCHLDSMRREKNAYYQD 61

Db 186 HPILFLRMMDRVKAGAKIIV--DPRRTATAEKADLFLQIAPGTDLALL----- 232

QY 62 GNALGVGVVSGTVEYDADRTYVVLQTDGRHTTSSFEHSPSR 106

Db 233 -NGLHLIVENGHT-----DREFIAETEGEVWPFLEQYTPDR 271

RESULT 8

QY 17 KHPLAQKRGARILVFTDDPRSRVLIVPGC-HLD---SMRREKNAYYQDGNALV-GMVVS 71

Db 122 REELASEAAAR-----ALAVPGCDHAEAELSVHRRTTHLLFADGDLAKESVQS 170

QY 72 GCTVEYDADD-----RTYVVQLTUDGRHTTSSFEHSPSRSP 108

Db 171 GGWRATADGSPGPGIRTYPER--TGRHAS-AGWEHIAALRLP 210

RESULT 9

ID Q9L6I6 PRELIMINARY; PRT; 483 AA.

AC Q9L6I6;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE LmbIH.

GN Name=lmbIH;

OS Streptomyces lincolnensis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI_TaxID=1915;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC25466;

RX MEDLINE=21591043; PubMed=11816970; DOI=10.1023/A:1012034421088;

RA Janata J., Najmanova L., Novotna J., Hola K., Felsborg J., Spizek J.;

RT "Putative lmbI and lmbH genes form a single lmbIH ORF in Streptomyces lincolnensis type strain ATCC 25466.";

RL Antonie Van Leeuwenhoek 79:277-284 (2001).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC25466;

RA Hola K., Janata J., Kopecky J., Novotna J., Najmanova L., Spizek J.;

RL Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AF234655; AAF65318.2; -

DR InterPro; IPR002510; Peptidase U62.

DR InterPro; IPR000408; Reg_chrom_condens.

DR Pfam; PF01523; EmdA_TlclD; 1.

DR PROSITE; PS00626; RCC1_2; UNKNOWN_1.

SQ SEQUENCE 483 AA; 50605 MW; 33F5AA07032944AC CRC64;

```

Query Match      12.5%; Score 74.5; DB 2; Length 483;
Best Local Similarity 32.0%; Pred. No. 31;
Matches 33; Conservative 9; Mismatches 36; Indels 25; Gaps 7;

QY 17 KHPAOKRGARILVFTDDPRRSVLIVPGC-HLD---SMREKNAYYFQDGNALV-GMVVS 71
Db 123 RELASEAAR-----ALAVPGCDHAEALSVRHRTLLIFADGGDLAKESVOS 171

QY 72 GGTVEYDADD-----RTYVQVLTQDGRHTTESFEHSSPSRSP 108
Db 172 GGVWRATADGPGSGIRTYP--TCRHAS-AGWEHIAALRLP 211

RESULT 10
ID Q9PFK0 PRELIMINARY; PRT; 576 AA.
AC Q9PFK0;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DE Alkaline phosphatase.
GN OrderedLocusNames=Xf0657;
OS Xylella fastidiosa
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN RP SEQUENCE FROM N.A.
RC STRAIN=9a5C;
RX MEDLINE=20365717; PubMed=10910347; DOI=10.1038/35018003;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Bartos M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Colombo A.A., Camargo L.E.A., Cartaro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.W., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
CC -!- SIMILARITY: Belongs to the alkaline phosphatase family.
DR EMBL; AS003910; AAF83467.1; -.
DR PIR; H82777; H82777.
DR HSSP; P05187; 1EW2.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001952; Alk_phosphatase.
DR Pfam; PF00245; Alk_phosphatase; 1.
DR PRINTS; PR00113; ALKPHPTASE.
DR SMART; SM00098; alkPPC; 1.
DR Complete proteome.
SQ SEQUENCE 576 AA; 61993 MW; 070FEC3184BF8B6D CRC64;

Query Match      12.4%; Score 73.5; DB 2; Length 576;
Best Local Similarity 28.3%; Pred. No. 48;

Matches 26; Conservative 7; Mismatches 56; Indels 3; Gaps 2;

QY 4 HAIRKHEMRQKPTKHP-LAQKRGARILVFTDDPRRSVLIVPGCHLDSMRREKNAYYFQD 62
Db 314 HMQPEHDDRHPNGEPSPSLAEMTRTAIQSLSRDPHGFLVMVEGGRIDHAHAGNAYRALDE 373

QY 63 NALVGMVVSGGTVEYDADDRTYVQVLTQDGRHT 94
Db 374 TIALSDAVRAAVQSAPKD--TLIIVTADHAHT 403

RESULT 11
MDR3_HUMAN
ID MDR3_HUMAN STANDARD; PRT; 1279 AA.
AC P21439;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Multidrug resistance protein 3 (P-glycoprotein 3).
GN Name=ABCB4; Synonyms=MDR3, PGY3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89138016; PubMed=2906314; DOI=10.1016/0378-1119(88)90057-1;
RA van der Blik A.M., Koolman P.M., Schneider C., Borst P.;
RT "Sequence of mdr3 cDNA encoding a human P-glycoprotein.";
RL Gene 71:401-411(1988).
RN [2]
RP SEQUENCE OF 856-1279 FROM N.A.
RX MEDLINE=88111519; PubMed=2892668;
RA van der Blik A.M., Baas F., ten Houte de Lange T., Koolman P.M.,
RA van der Velde-Koerts T., Borst P.;
RT "The human mdr3 gene encodes a novel P-glycoprotein homologue and
RT gives rise to alternatively spliced mRNAs in liver.";
RL EMBO J. 6:3325-3331(1987).
RN [3]
RP GENE STRUCTURE.
RX MEDLINE=91161629; PubMed=2002063;
RA Lincke C.R., Smit J.J.M., van der Velde-Koerts T., Borst P.;
RT "Structure of the human MDR3 gene and physical mapping of the human
RT MDR locus.";
RL J. Biol. Chem. 266:5303-5310(1991).
CC -!- FUNCTION: Energy-dependent efflux pump responsible for decreased
CC drug accumulation in multidrug-resistant cells. Human MDR3 is not
CC capable of conferring drug resistance. Mediates the translocation
CC of phosphatidylcholine across the canalicular membrane of the
CC hepatocyte.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- DISEASE: Defects in ABCB4 are the cause of progressive familial
CC intrahepatic cholestasis type III (PFIC3) [MIM:602347]. PFIC3 is
CC an autosomal recessive liver disorder presenting with early onset
CC cholestasis that progresses to cirrhosis and liver failure before
CC adulthood. It is characterized by elevated serum gamma-
CC glutamyltransferase levels.
CC -!- SIMILARITY: Belongs to the ABC transporter family. MDR subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
DR EMBL; M23234; AAA36207.1; -.
DR EMBL; X06181; CAA29547.1; -.
DR PIR; JS0051; DVHU3.
DR HSSP; P08716; 1MT0.
DR Genew; HGNC:45; ABCB4.
DR MIM; 171060; -.

```

DR MIM; 602347; -.
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0005624; C:membrane fraction; TAS.
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; TAS.
 DR GO; GO:0042493; P:response to drug; TAS.
 DR InterPro; IPRO03593; AAA_ATPase.
 DR InterPro; IPRO01140; ABC_transport.
 DR InterPro; IPRO03439; ABC_TW_transport.
 DR Pfam; PF00664; ABC membrane; 2.
 DR Pfam; PF00005; ABC tran; 2.
 DR ProDom; PD000006; ABC transporter; 2.
 DR SMART; SM00382; AAA; 2.
 DR PROSITE; PS00929; ABC_TMIF; 2.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
 DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
 KW ATP-binding; Glycoprotein; Multigene family; Repeat; Transmembrane;
 KW Transport.
 FT DOMAIN 1 57 Cytoplasmic (Potential).
 FT TRANSMEM 58 78 Potential.
 FT TRANSMEM 123 143 Potential.
 FT TRANSMEM 192 211 Potential.
 FT TRANSMEM 216 235 Potential.
 FT TRANSMEM 301 320 Potential.
 FT TRANSMEM 336 354 Potential.
 FT DOMAIN 355 711 Cytoplasmic (Potential).
 FT TRANSMEM 712 732 Potential.
 FT TRANSMEM 756 776 Potential.
 FT TRANSMEM 832 851 Potential.
 FT TRANSMEM 854 873 Potential.
 FT TRANSMEM 937 956 Potential.
 FT TRANSMEM 976 993 Potential.
 FT DOMAIN 994 1279 Cytoplasmic (Potential).
 FT NP_BIND 429 436 ATP (By similarity).
 FT NP_BIND 1069 1076 ATP (By similarity).
 FT REPEAT 1 640
 FT REPEAT 641 1279
 FT CARBOHYD 91 91 N-linked (GlcNAc. .) (Potential).
 FT CARBOHYD 97 97 N-linked (GlcNAc. .) (Potential).
 FT CONFLICT 1093 1093 V -> VFVDFGFQ (in Ref. 2).
 FT SEQUENCE 1279 AA; 140682 MW; 3D58C98B5C8D5087 CRC64;

 Query Match 12.4%; Score 73.5; DB 1; Length 1279;
 Best Local Similarity 24.2%; Pred. No. 1.2e+02;
 Matches 23; Conservative 21; Mismatches 32; Indels 19; Gaps 5

 QY 18 HPLAQKRGARILVFTDDPRRSVLIVPGCHLDSMRKNAYVFDGNALVGMVVGQTV-- 75
 Db 589 HRLSTVRNADVTAGFED--GVIVEQGHSHSLMKKE--GVYFK----LVNMQTSGSQIQS 639

 QY 76 -EYDADRTYVVQLTDG-----RHTTSSFEHS 102
 Db 640 EFPELNDEKAATRMAPNGWKSFLFRHSQTKNLKN 674

 RESULT 12
 Q8MHZ3 PRELIMINARY; PRT; 215 AA.
 AC Q8MHZ3
 DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
 DE SVS (Fragment).
 GN Name-SRV.
 OS Callithrix aurita (White-eared marmoset) (Buffy-eared marmoset).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Callithrix.
 ON NCBI_TaxID=57375;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood.
 RX MEDLINE=22155426; PubMed=12165946; DOI=10.1007/s00299-001-2308-7;
 RA Moreira M.A.M.;
 RT "SRV evolution in Cebidae (Platyrrhini: Primates).";

RESULT 14

Q65KF1 PRELIMINARY; PRT; 516 AA.
 AC Q65KF1;
 DT 25-OCT-2004 (T-EMBLrel. 28, Created)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
 DE P6SX (Hypothetical protein).
 GN Names=P6SX; ORFNames=BL00868, BL101563;
 OS Bacillus licheniformis DSM 13.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=279010;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 13;
 RX PubMed=15383718;
 RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
 RA Ehrenreich P., Baeumer S., Henne A., Liesegang H., Merkl R.,
 RA Ehrenreich A., Gottschalk G.;
 RT "The Complete Genome Sequence of Bacillus licheniformis DSM13, an
 RT Organism with Great Industrial Potential";
 RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 14580;
 RA Rey M.W., Ramaiya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
 RA Tang M., de Leon A.L., Xiang H., Gusti V., Clausen I.G., Olsen P.B.,
 RA Rasmussen M.D., Andersen J.T., Jorgensen P.L., Larsen T.S.,
 RA Sorokin A., Bolotin A., Lapidus A., Galleron N., Ehrlich S.D.,
 RA Berka R.M.;
 RT "Complete genome sequence of the industrial bacterium Bacillus
 RT licheniformis and comparisons with closely related Bacillus species";
 RL Genome Biol. 5:R77-R77(2004).
 DR EMBL; AB017333; AAU40463.1;
 DR EMBL; CP000002; AAU23108.1;
 KW Hypothetical protein.
 SQ SEQUENCE 516 AA; 57348 MW; 3FA5E604AA15B7BC CRC64;

Query Match 12.2%; Score 72.5; DB 2; Length 516;
 Best Local Similarity 22.4%; Pred. No. 54;
 Matches 22; Conservative 21; Mismatches 38; Indels 17; Gaps 4;
 QY 16 TKHPLAQR-----GARILVFTDPRRSVLIVPGCHLDSMRREKNAYYFQGNALVGMVVS 71
 DB 267 TRHITADERTKWDGGLRLKLTDDSGKRTQVADGADLLTL---STGFYFATGNVQNNPTT 323
 QY 72 GGTVEYDADRTYVVLTDGCR-----HTTESSFEHSS 103
 DB 324 NDSAQFTYD-----VIESDQGRKTIYAWRTIDTLWHTAT 357

RESULT 15

Q6CKY2 PRELIMINARY; PRT; 546 AA.
 AC Q6CKY2;
 DT 25-OCT-2004 (T-EMBLrel. 28, Created)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
 DE Similarities with sgd|S0002247 Saccharomyces cerevisiae YDL089w.
 GN ORFNames=KLLA0F07205g;
 OS Kluyveromyces lactis NRRL Y-1140.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
 OX NCBI_TaxID=284590;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL Y-1140;
 RG Genolevures;
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
 RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
 RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
 RA Boisrame A., Boyer J., Cattolico L., Confanieri F., de Daruvar A.,

RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
 RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
 RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
 RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
 RA Swennene D., Tekalia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
 RA Zentou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
 RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
 RT Wincker P., Souciet J.L.;
 RT "Genome evolution in Yeasts";
 RL Nature 430:35-44(2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL Y-1140;
 RA Genoscope;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBDJ databases.
 DR EMBL; CR382126; CAG98115.1;
 DR InterPro; IPR002048; EF-hand.
 DR PROSITE; PS00018; EF_HAND; UNKNOWN 1.
 SQ SEQUENCE 546 AA; 64086 MW; DC7F07417145E331 CRC64;
 Query Match 12.2%; Score 72.5; DB 2; Length 546;
 Best Local Similarity 23.6%; Pred. No. 57;
 Matches 25; Conservative 21; Mismatches 43; Indels 17; Gaps 4;
 QY 4 HAIRKH-MRQRKPTKHPKLAQKRGARILVFTDPRRSVLIVPGCHL---DSMRREKNAYYF 59
 DB 451 NASRHNHISERRNSHSPKQHEANRLNYPDETNDISPDVSDRFRMDDRFRGRGQYFN 510
 QY 60 QDGNALVGMVSGGTVEYDADRTYVVLTDGCRHTTESSFEHSSPS 105
 DB 511 RSPD-----INSGTLHYDDGD-----DDNDRISKSPFRNSSSS 543

Search completed: October 15, 2005, 04:35:52
 Job time : 26.3113 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:20:35 ; Search time 101.02 Seconds
(without alignments)
1454.845 Million cell updates/sec

Title: US-10-617-038-6
Perfect score: 1991
Sequence: 1 VAGNPDVTVLLGGDVMLGR.....VTRPDNLLEVPAAMLTSE 380

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1991	100.0	380	ADI37285	Adi37285 M. tuberc
2	252	12.7	380	AAB74026	Aab74026 Bacillus
3	194	9.7	323	ABM67261	Abm67261 Photorhab
4	183	9.2	372	ABBA49822	Abba49822 Listeria
5	182.5	9.2	413	AAU63427	Aau63427 Propionib
6	182.5	9.2	413	ABMS9946	Abms9946 Propionib
7	179	9.0	398	ADB06120	Adb06120 Alloococ
8	177	8.9	476	ABBA47908	Abba47908 Listeria
9	172	8.6	322	ADB06118	Adb06118 Alloococ
10	169.5	8.5	594	ADA34203	Ada34203 Acinetoba
11	165	8.3	447	ABP66202	Abp66202 Bifidobac
12	153.5	7.7	249	ADB06116	Adb06116 Alloococ
13	150	7.5	469	ADB08332	Adb08332 Alloococ
14	150	7.5	474	ADB08334	Adb08334 Alloococ
15	137.5	6.9	257	ABP79262	Abp79262 N. gonorr
16	128	6.4	666	AAAB79657	Aab79657 Corynebac
17	128	6.4	666	AAU71885	Aau71885 C. glutam
18	128	6.4	666	ADD13553	Add13553 C. glutam
19	125.5	6.3	201	ADR94602	Adr94602 Novel S.
20	115.5	5.8	170	ABU01903	Abu01903 S. pneumo
21	112.5	5.7	416	ABP29972	Abp29972 Streptoco
22	112.5	5.7	430	ABP26343	Abp26343 Streptoco
23	112	5.6	961	ADA36071	Ada36071 Acinetoba
24	108.5	5.4	3546	AAAB82212	Aab82212 Polyketid
25	107.5	5.4	4861	AAU84280	Aau84280 Human end

ALIGNMENTS

RESULT 1

ADI37285

ID ADI37285 standard; protein; 380 AA.

XX AC ADI37285;

XX DT 22-APR-2004 (first entry)

DE M. tuberculosis low oxygen induced antigen Rv0574c SEQ ID NO:6.

XX mycobacterial infection; vaccine; tuberculosis;

KW Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;

KW low oxygen induced antigen.

OS Mycobacterium tuberculosis.

XX WO2004006952-A2.

XX PD 22-JAN-2004.

XX PF 08-JUL-2003; 2003WO-DK000477.

XX PR 13-JUL-2002; 2002DK-00001098.

XX (STAT-) STATENS SERUM INST.

XX Andersen P, Rosenkrands I, Stryhn A;

XX WPI; 2004-122778/12.

XX N-PSDB; ADI37330.

Use of one or more polypeptides or their fragments, which are expressed during the latent stage of the mycobacterial infection, and/or nucleic acids encoding the polypeptides, for a therapeutic vaccine against tuberculosis.

Claim 3; SEQ ID NO 6; 76pp; English.

The present invention describes polypeptides or their fragments, which are expressed during the latent stage of a mycobacterial infection, and/or nucleic acids encoding the polypeptides, which are useful for creating a therapeutic vaccine against tuberculosis. Also described: (1) a therapeutic vaccine against tuberculosis comprising one or more polypeptides; (2) a method for treating an animal, including a human being, with tuberculosis caused by virulent mycobacteria, e.g. by Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for immunising an animal, including a human being, against tuberculosis caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis

CC caused by virulent mycobacteria in an animal, including a human being;
CC (5) a method for diagnosing previous or ongoing infection with a virulent
CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
CC infection in a subject. The polypeptides have antibacterial activities,
CC and can be used in vaccines and in gene therapy. The polypeptides are
CC useful for the manufacture of a therapeutic vaccine for treating an
CC individual who is infected by a virulent mycobacterium, e.g. M.
CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
CC The present sequence represents a low oxygen induced antigen, which is
CC used in the exemplification of the present invention.

XX
SQ Sequence 380 AA;

Query Match 100.0%; Score 1991; DB 8; Length 380;
Best Local Similarity 100.0%; Pred. No. 4.8e-190; Mismatches 0; Indels 0; Gaps 0;
Matches 380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAGNPDVVTVLLGGDVMGLRGVDQILPHPGKPOLRERYMRDATGYVRLAERVNGRIPLPV 60
DB 1 VAGNPDVVTVLLGGDVMGLRGVDQILPHPGKPOLRERYMRDATGYVRLAERVNGRIPLPV 60
QY 61 DWRWPWGALAVLENTATDVCLINLETTITADGEFADRPKPCYRMHPDNVPALTALRPHV 120
DB 61 DWRWPWGALAVLENTATDVCLINLETTITADGEFADRPKPCYRMHPDNVPALTALRPHV 120
QY 121 CALANNHILDFCYQGLTDTVAALACAGIQSVGAGADLLAARRSALVTYGHERRVIVGSA 180
DB 121 CALANNHILDFCYQGLTDTVAALACAGIQSVGAGADLLAARRSALVTYGHERRVIVGSA 180
QY 181 AESSGVPSWAARRDRPGVWLIIRDPAQRDVAADVAQVLADKRPDGIIVSNMHWGSNWG 240
DB 181 AESSGVPSWAARRDRPGVWLIIRDPAQRDVAADVAQVLADKRPDGIIVSNMHWGSNWG 240
QY 241 ATPAGDVAFAHRLIDAGIDMVHGHSSHPRIEYIRGKPILYCGDGVDDYEGIGGHSF 300
DB 241 ATPAGDVAFAHRLIDAGIDMVHGHSSHPRIEYIRGKPILYCGDGVDDYEGIGGHSF 300
QY 301 RSELRLLYLTVTPDASGNLISIQMLPLRVSRMLQRASQDTDEWLNTIERSRFGIRV 360
DB 301 RSELRLLYLTVTPDASGNLISIQMLPLRVSRMLQRASQDTDEWLNTIERSRFGIRV 360
QY 361 VTRPDNLLEVPAAALTSKE 380
DB 361 VTRPDNLLEVPAAALTSKE 380

RESULT 2
AAB74026
ID AAB74026 standard; protein; 380 AA.

AC AAB74026;

XX 21-JUN-2001 (first entry)

DE Bacillus subtilis IFO 3336 PGA synthesising enzyme.

KW Bacterial; poly-gamma-L-glutamic acid; L-PGA;

KW poly-gamma-D,L-glutamic acid; PGA; glutamate racemase.

XX Bacillus subtilis.

OS JP2001017182-A.

PN 23-JAN-2001.

XX 09-JUL-1999; 99JP-00196335.

XX 09-JUL-1999; 99JP-00196335.

XX (NAGS) NAGASE SANGYO KK.

XX WPI; 2001-285408/30.

DR N-PSDB; AAF82257.

XX New nucleic acid encoding a glutamate racemase enzyme useful for the
PT preparation of poly-gamma-glutamic acid.

XX Disclosure; Page 13-15; 17pp; Japanese.

XX The present sequence is an enzyme which is useful in the production of
CC poly-gamma-L-glutamic acid (L-PGA) or poly-gamma-D,L-glutamic acid (PGA).
CC A plasmid comprising the present sequence may be used to transform
CC Escherichia coli. The transformants express the enzyme and PGA is
CC produced in the culture

XX Sequence 380 AA;

Query Match 12.7%; Score 252; DB 4; Length 380;

Best Local Similarity 26.0%; Pred. No. 4.6e-16;

Matches 77; Conservative 51; Mismatches 104; Indels 64; Gaps 11;

QY 6 DWVTVLLGGDVMGLRGVDQILPHPGKPOLRERYMRDATGYV----RLAERYNGRIPLPVD 61
DB 57 DVLASFPVGDIMRGYVEKVTQKQADSIFQ-----YVEPIFRASDYVAGNFENPVT 108
QY 62 WRWPWGALAVLENTATDVCLINLETTITADGEFADRPKPCYRMHPDNVPALTALRPHVC 121
DB 109 YQNVYKQA-----DKEIHLQTNKESVKVLKDMFTVL 140
QY 122 ALANNHILDFCYQGLTDTVAALACAGIQSVGAGADLLAARRSALVTYGHERRVIVGSVAA 181
DB 141 NSANNHAMDYGVQGMKDTLGEFAKQNLIDIVGAGYSLSDAKKK---ISYQK---VNGVTI 193
QY 182 ESSGVPE----SWAARRDRPGVWLIIRDPAQRDVAADVAQVLAD-KRPDGIIVSNMHWGS 236
DB 194 ATLFTDVSXGKGFPAKKNTPCV-LPADP-----EIFPMISEAKKHADIIVVQSHWGO 245
QY 237 NWGATAPGD--VAFARHLIDAGIDMVHGHSSHPRIEYIRGKPILYCGDGVDD 290
DB 246 E--YDNDPNDQRQLARAMSADAGDIIVGHHPVLEPIEVYNGTVIFYSLGNFVFD 299

RESULT 3

ABM67261

ID ABM67261 standard; protein; 323 AA.

AC ABM67261;

XX 20-NOV-2003 (first entry)

DE Photorhabdus luminescens protein sequence #358.

KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
KW detection; food; gene expression; plant; animal; microorganism; toxin;
KW antibiotic; biopesticide; virulence factor; disease model; plague;
KW whooping cough.

OS Photorhabdus luminescens.

PN WO200294867-A2.

XX 28-NOV-2002.

XX 07-FEB-2002; 2002WO-IB003040.

XX 07-FEB-2001; 2001FR-00001659.

XX (INSP) INST PASTEUR.

PA (CNRS) CNRS CENT NAT RECH SCI.

XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;

XX Buchrieser C;

XX WPI; 2003-148459/14.

XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,

CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a specifically claimed polypeptide which is predicted
CC to be encoded by an ORF (open reading frame) contained within the P.
CC acnes polynucleotides of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 413 AA;

Query Match 9.2%; Score 182.5; DB 6; Length 413;
Best Local Similarity 23.8%; Pred. No. 4.7e-09;
Matches 89; Conservative 51; Mismatches 149; Indels 85; Gaps 17;

QY 42 ATGYVRLAERVNGRIPLV--DWRW---PWGEA-----LAVLENTATD 79
Db 70 ATTQPIHRTNTGSLTVTVSGDLLMHPSTWGTAREDGKNDPAPITGTVAPILRN--AD 127
QY 80 VCLINLETTITADGEFADRPKPCYRMHPD-NVPA-----LTALRPHVCALANNHILDFGY 133
Db 128 VSIICHEVPVAPKGS-----QYSGYPEFAPVPAETAKGIAVGFDPACSTASHSFDRL 180
QY 134 QGLTDTVAALAGAGIOSVGAGADLLAARSALVTVGHERRVIVGSVAESSGVP--ESWA 191
Db 181 PGVKATLDLAAAHVHSGTARTKEADTPVIVSHGLKGLVSGAYGLNGSTPPKGSWA 240
QY 192 ARREDRGWMLRDPQARDV-ADDVAQVLADKEPG-DIAIVSMHWSNMGYATAPGDVAF 249
Db 241 -----W-----SDIEADHLIKRAEAKKAGADIVIAAHSGLEYHHPTGQIRL 285
QY 250 AHRL-IDAGIDVMHGHSHHPRPIETIRGPKILYGGCDVVDDY-----EGIGGHEGF 300
Db 286 AQELTASPAVDVYCHSHVVEPTWRMNGKIWMYGLNLVQAQSSNMPGTDEGVGRVTF 345
QY 301 R-----SELRLYLTV-----TDPASGNLISLQMLPLRVSRMRLQRASQTDTEWLRNT 348
Db 346 TVRSGKVTTKAEYIPILIGSKNDGPRIHAVDTLKSNGMGNQARLKSQAQ-----D 397
QY 349 IERISRRFGIRVVT 362
Db 398 ISRIVTSLGVDGVT 411

RESULT 7
ADB06120
ID ADB06120 standard; protein; 398 AA.
XX
AC ADB06120;
XX
DT 20-NOV-2003 (first entry)
XX
DE Alloiococcus otitis antigenic protein SEQ ID NO:60.
XX
KW Alloiococcus otitis; antigenic protein; immunogenic; immunisation;
KW gene therapy; Gram-positive bacterium; infection.
XX
OS Alloiococcus otitis.
XX
PN W02003048304-A2.
XX
PD 12-JUN-2003.
XX
PF 25-NOV-2002; 2002WO-US036123.
XX

PR 29-NOV-2001; 2001US-0333777P.
PR 18-NOV-2002; 2002US-0426742P.
XX
PA (AMHP) WYETH HOLDINGS CORP.
XX
PI Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;
XX
DR WPI; 2003-505284/47.
DR N-PSDB; ADB06119.
XX
XX New Alloiococcus otitis polynucleotides and polypeptides, useful for
XX treating and diagnosing diseases, drug screening assays and monitoring of
XX effects during drug clinical trials.
PS
PS Claim 33; SEQ ID NO 60; 1019pp; English.
XX
XX The present invention describes an isolated polynucleotide (I) of
XX Alloiococcus otitis genomic DNA, which encodes an antigenic protein.
XX Alloiococcus otitis is a Gram-positive bacterium. Also described: (1)
XX an isolated polypeptide that is encoded by the polynucleotide (I); (2) an
XX expression vector comprising the novel isolated polynucleotide (I), its
XX complement, degenerate variant or fragment; (3) a genetically engineered
XX host cell, transfected, transformed or infected with the vector of (2);
XX (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
XX composition comprising the polypeptide, its complement, biological
XX equivalent or fragment, or the polynucleotide that is comprised in the
XX expression vector; (6) a pharmaceutical composition comprising the
XX polypeptide of (1) and a carrier; (7) a protein chip comprising an array
XX of the polypeptides of (1), their biological equivalent or fragment; (8)
XX immunising against Alloiococcus otitis by administering to a host the
XX immunogenic composition; (9) detecting and/or identifying Alloiococcus
XX otitis in the biological sample; (10) a kit comprising a container
XX containing the novel polynucleotide, its degenerate variant or fragment,
XX or the antibody of (4); and (11) producing a polypeptide by culturing the
XX genetically engineered host cell under conditions suitable to produce the
XX polypeptide from the culture. (I) can be used in gene therapy. The
XX polynucleotides, polypeptides, antibodies and compositions of the present
XX invention can be used for treating and diagnosing diseases, drug
XX screening assays and monitoring of effects during drug clinical trials.
XX The polynucleotides are useful for expressing and detecting Alloiococcus
XX otitis. The present sequence represents an Alloiococcus otitis
XX antigen protein from the present invention.
SQ Sequence 398 AA;

Query Match 9.0%; Score 179; DB 6; Length 398;
Best Local Similarity 20.9%; Pred. No. 9.9e-09;
Matches 83; Conservative 61; Mismatches 163; Indels 90; Gaps 11;

QY 1 VAGNPDVTVTLGGDVMILGRGVDQILPHPGKPOLRERYMRDATGYVRLAERVNGRIPLV 60
Db 65 VQANAETPSISFVGDISLSRYIAEIGERDGYEEF-----FTYVKPHFDGR----- 109
QY 61 DMRWPGALAVLENTATDVCLINLETTITAD-----GEFADRPKPCYRMHP 107
Db 110 -----DLVVANLESATVYDNRVTYVKPASSLSGGIYLDSS-----I 144
QY 108 DNVLPALTALRPHVCALANNHILDFGVGLTDTVAALAGAGIOSVGAGADLLAARSALVT 167
Db 145 DSVKAMQAGIDLVSMMANNHTGDMKQGVMDGMEILRDVEMDYIGHGHDRAEQAYQFV 204
QY 168 VGHERRVIVGSVAESSGVPESWAARRDRPGVWMLIRDPAQRDVAADDVAQVLADKRPGDI 227
Db 205 ANN---LTTTSIFAVSDVIKPGQAASDDPEGLVLTTSNQAFNLANSY-----DQESDL 253
QY 228 AIVSMHWSNMGYATAPGDVAFARHLIDAGIDMVHGHSHHPRPIETIRGPKILYCCGDV 287
Db 254 VIAYIHAGIETIRQPDANHOELAESLIDAGADIVICSHTHSLPLPEKYQDGIIFYGLGNP 313
QY 288 VDDYEGIGGHEFSRSELRLYLTVTDPASGNLISLQMLPLRV-----SRMRLQRAS 338
Db 314 IFDQ-----GMQSSDTSV-ILDMDINSP---DQVRFTLRPMKIEAGIPKPNGLVTRIQ 365

CC containing the novel polynucleotide, its degenerate variant or fragment,
CC or the antibody of (4); and (11) producing a polypeptide by culturing the
CC genetically engineered host cell under conditions suitable to produce the
CC polypeptide from the culture. (1) can be used in gene therapy. The
CC polynucleotides, polypeptides, antibodies and compositions of the present
CC invention can be used for treating and diagnosing diseases, drug
CC screening assays and monitoring of effects during drug clinical trials.
CC The polynucleotides are useful for expressing and detecting *Alloicoccus*
CC otitidis. The present sequence represents an *Alloicoccus* otitidis
CC antigen protein from the present invention.
XX
XX
SQ Sequence 322 AA;

Query Match 8.6%; Score 172; DB 6; Length 322;
Best Local Similarity 21.1%; Pred. No. 3.7e-08;
Matches 81; Conservative 58; Mismatches 155; Indels 90; Gaps 11;

QY 14 GDVMLGRGVDDQLPHKPKQLBRYNRDGTGYVRLAERYNGRIPLPVDWRWPGMALAVL 73
DB 2 GDLSLSRYIAEIGERDGYBEF-----FTYVKPHFDGR----- 33

QY 74 ENTATDVCLINLETTITAD-----GEFADRPVCVYRMHPDNPVLTALRPHV 120
DB 34 -----DLVVANLESATYDNRVTYKVPASSLSGGIYLDSS-----IDSVKAMQEGIDL 81

QY 121 CALANNHILDYFGQGLTDTVAALAGAGIOSVGAGADLLAARRSALVTVGHERRVIVGSA 180
DB 82 VSMANNHTGDMGQGVMDGMEILRDVEMDYIGMHDRAEAGQAYQVANN-----LTTTSIF 137

QY 181 AESSGVPESSAARRDRPGVWLIRDPQARDVADDDVAQVLADKRPDGIATVSMHWSNWCY 240
DB 138 AVSDVTKPQOASDDDFGLTTSNQAFNLANSY-----DOESDLVIAYIHAGIEYIR 190

QY 241 ATAPGDVAFARHLIDAGIDMVHSHHPRPIEYRGKPILYCGGVDDYDEGIGHESF 300
DB 191 QPDANHQELAESLIDAGADIVICSHTHSLLPVEKYQDGLIFYGLGNFIFDQ-----GMQSS 246

QY 301 RSELRLLYTVPDASGNLISQMLPLRV-----SRMLQRASQDTTEWLRNTIER 351
DB 247 TDSV-ILDMDINSF---DQVRFTRPMKIEAGIPKPNGLVTRIQIKRLTERLDQDFK 302

QY 352 ISRRFGIRVTRPDNLEVVPAAN 375
DB 303 ID-----RADNVIDFDLAN 317

RESULT 10
ADA34203
ID ADA34203 standard; protein; 594 AA.
AC ADA34203;
XX
XX
DT 20-NOV-2003 (first entry)
XX
XX
DE Acinetobacter baumannii protein #1364.
XX
XX
KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
KW plant biocontrol agent.
XX
XX
OS Acinetobacter baumannii.
XX
PN US6562958-B1.
XX
PD 13-MAY-2003.
XX
PF 04-JUN-1999; 99US-00328352.
XX
PR 09-JUN-1998; 98US-0088701P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton G, Bush D;
XX

DR WPI; 2003-576092/54.
DR N-PSDB; ADA30077.
XX
PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT for diagnosing a bacterial disease, as components of antibacterial
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT plants.
XX
PS Example; SEQ ID NO 5490; 328pp; English.
XX
CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC The A. baumannii nucleic acids and polypeptides are useful as reagents
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of
CC A. baumannii and other Acinetobacter species in a sample, in screening
CC compounds for the ability to interfere with the A. baumannii life cycle
CC or to inhibit A. baumannii infection, and as biocontrol agents for
CC plants. The present sequence represents the amino acid sequence of an A.
CC baumannii protein.
XX
SQ Sequence 594 AA;

Query Match 8.5%; Score 169.5; DB 6; Length 594;
Best Local Similarity 25.3%; Pred. No. 1.6e-07;
Matches 56; Conservative 41; Mismatches 109; Indels 15; Gaps 6;

QY 79 DVCLINLETTITADGE--FADRKVCYRMH-PDNPALTALRPHVCALANNHILDFGYQG 135
DB 305 DINIANFEAVFSLNQSPKDKKPFILKADAKKTLEEFKSIHLNVLVLANNHLKDYEGQ 364

QY 136 LTDVVAALAGIOSVGAGADLLAARRSALVTVGHERRVIV-GSVAESSGVPESSAARR 194
DB 365 LAYTLQLDQASISYIGAGLNQKDAHNYFEITFTFKHYAIFNGYWHRDYALDYDFYALG 424

QY 195 DRPGVWLIRDPQARDVADDDVAQVLADK--PGDIAIVSMHWSNGVATAPGDVAFAR 252
DB 425 SRSGVACLVG-----LLEQIMRYKQAHPEKLIIVICHWGVDFKPI--KDQTKLATI 475

QY 253 LIDAGIDMVHSHHPRPIEYRGKPILYCGGVDDYDEG 293
DB 476 LTQAGADLVHGHGANTIQIPIQIINQKPVVFNIGNAVFNSDG 516

RESULT 11
ABP66202
ID ABP66202 standard; protein; 447 AA.
AC ABP66202;
XX
XX
DT 19-NOV-2002 (first entry)
XX
DE Bifidobacterium longum NCC2705 ORF amino acid sequence SEQ ID NO:946.
XX
KW Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
KW antidiarrheic; antibacterial; inhibitor of Salmonella; detection;
KW identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
KW rotavirus; food composition; pharmaceutical composition.
XX
OS Bifidobacterium longum.
XX
PN EP1227152-A1.
XX
PD 31-JUL-2002.
XX
PF 30-JAN-2001; 2001EP-00102050.
XX
PR 30-JAN-2001; 2001EP-00102050.
XX
PA (NEST) SOC PROD NESTLE SA.
XX
XX WPI; 2002-668397/72.
XX
PT Novel polynucleotide comprising Bifidobacterium genome sequence useful as

PT a probe or primer for detecting and/or identifying *Bifidobacterium longum*
PT in a biological sample.

PS Claim 3; SEQ ID NO 946; 80pp; English.

The present invention describes a polynucleotide (I) comprising a sequence of a Bifidobacterium genome selected from the nucleotide sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at least 90% identity or which hybridises with the sequences given in ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding a fusion protein, comprising a sequence selected from 1097 sequences given in APB65258 to APB65354 ligated in frame to a polynucleotide encoding a heterologous polypeptide. (I) has antidiarrheic and antibacterial activities, and can be used as an inhibitor of *Salmonella*. (II) (which is a probe) is useful for the detection and/or identification of *Bifidobacterium longum* in a biological sample. A carrier containing the lactic acid bacterium *Bifidobacterium longum* NCC3705 (GNCM 1-2618) can be used for preventing and/or treating diarrhoea brought about by pathogenic bacteria and/or rotavirus. The carrier is a food composition selected from milk, yogurt, curd, cheeses, fermented milks, milk based fermented products, ice-creams, fermented cereal based products, milk based powders, infant formula, pet food or a pharmaceutical composition selected from tablets, liquid bacterial suspensions, dried oral supplement, wet oral supplement, dry tube feeding or wet tube feeding. (I) is useful in DNA arrays or chips to carry out analysis of the expression of the *Bifidobacterium* gene. ABQ81844 to ABQ81850 represent *Bifidobacterium* related nucleotide sequences given in the Sequence Listing from the present invention but not mentioned further within the specification. N.B. The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied by the European Patent Office

Sequence 447 AA;

Query Match	8.3%;	Score 165;	DB 5;	Length 447;
Best Local Similarity	23.1%;	Pred. No. 3e-07;		
Matches	53;	Conservative 40;	Mismatches 106;	Indels 30; Gaps 7;
Qy	77	ATDVCLINLETTITADGEFAEDRKPVCYRMHPDNVPALTALRPHVCAIANNHILDFGVQGL	136	
Dd	132	ASDIAVCEFEPTIAQRGGPYAAYPI-FNIPPEVADAARNVGYNACTHATNSHWPGQADGI	190	
Qy	137	TDTVAALLAGAGIQSGAGADLLAARRSALV---TVGHERRIVVIGSVAAESSGVPESWAAR	193	
Dd	191	ARLWDTLEANGIAQTGSVKTEADSTKLPIVIDSPTGCGKGLVTGTSLNGMTADEDKQVD	250	
Qy	194	RDRPGVWLIRDPARDVDVVAAOV----LADKRPGDIAIVSMI-----WGSNWGYAT	242	
Dd	251	R-----LRDDGPNDHSIDTQRAVDKANKAREOGADVVALMHSSVOBYLDYADSW----	299	

RESULT 12

ADB06116

ID ADB06116 standard; protein; 249 AA.

XX

AC ADB06116;

XX
DT 20-NOV-2003 (first entry)

XX DE Alloiococcus otitis antigenic protein SEQ ID NO:56.

xx
xx
KW Alloicoccus otitidis; antigenic protein; immunogenic; immunisation;
KW gene therapy; Gram-positive bacterium; infection.

XX OS **Alloiococcus otitis.**

XX
PN WO2003048304-A2.

XX
PD 12-JUN-2003.

XX	25-NOV-2002; 2002WO-US036123.
PF	
XX	
XX	29-NOV-2001; 2001US-0333777P.
PR	
PR	18-NOV-2002; 2002US-0426742P.
XX	
XX	(AMHP) WYETH HOLDINGS CORP.
PA	
XX	
XX	Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;
PI	
XX	WPI; 2003-505284/47.
DR	N-PSDB; ADB06115.
DR	
XX	
XX	New Alloicoccus otitis polynucleotides and polypeptides, useful for
PT	treating and diagnosing diseases, drug screening assays and monitoring of
PT	effects during drug clinical trials.
PT	
XX	Claim 33; SEQ ID NO 56; 1019pp; English.
XX	
PS	

The present invention describes an isolated polynucleotide (1) of *Alloccoccus* otitidis genomic DNA, which encodes an antigenic protein. *Alloccoccus* otitidis is a Gram-positive bacterium. Also described: (1) an isolated polypeptide that is encoded by the polynucleotide (1); (2) an expression vector comprising the novel isolated polynucleotide (1), its complement, degenerate variant or fragment; (3) a genetically engineered host cell, transfected, transformed or infected with the vector of (2); (4) an antibody specific for the polypeptide of (1); (5) an immunogenic composition comprising the polypeptide, its complement, biological equivalent or fragment, or the polynucleotide that is comprised in the expression vector; (6) a pharmaceutical composition comprising the polypeptide of (1) and a carrier; (7) a protein chip comprising an array of the polypeptides of (1), their biological equivalent or fragment; (8) immunising against *Alloccoccus* otitidis by administering to a host the immunogenic composition; (9) detecting and/or identifying *Alloccoccus* otitidis in the biological sample; (10) a kit comprising a container containing the novel polynucleotide, its degenerate variant or fragment, or the antibody of (4); and (11) producing a polypeptide by culturing the genetically engineered host cell under conditions suitable to produce the polypeptide from the culture. (1) can be used in gene therapy. The polynucleotides, polypeptides, antibodies and compositions of the present invention can be used for treating and diagnosing diseases, drug screening assays and monitoring of effects during drug clinical trials. The polynucleotides are useful for expressing and detecting *Alloccoccus* otitidis. The present sequence represents an *Alloccoccus* otitidis antigen protein from the present invention.

Sequence 249 AA;

Query Match	7.7%;	Score 153.5;	DB 6;	Length 249;
Best Local Similarity	23.6%;	Pred. No. 1.8e-06;		

QY 122 ALANNHILDFGYOGLTDTVAALAGAGIOSVGAGADLLAARRSALVTVGHERRVI VGSVAA 181

Db 10 SMANNHTGDMKQGVMDGMEILRDVEMDYIGMHDPRAEAGAYQFVANN ---LTT SIFA 65

QY 182 ESSGVPESWAARRDRPGVWLI RDPAQRDVADDVAAQVLADKRPGDI AIVSMHWGCSNWGYA 241

Db 66 VSDVIKPGQAAASDDDEPGVLTNTNSQAFLNLSY-----DOESDLVIAIYHAGIEYIRO 118

Qy 242 TAPGDVAFARLIDAGIDMVHGSHHPRPIEIRGKPILYCGDVVDDEGIGGHESFR 301

[illegible]

QY 302 SELRLLYLTVTDPASGNLISLQMLPLRV-----SRMRLQASQTDTEWLNRNTIERI 352

[illegible]

QY 353 SRRFGIRVTRPDNLLLEVPAAN 375

Db 231 D-----RADNIVIDFLAN 244

```

RESULT 13
ADB08322 ID ADB08322 standard; protein; 469 AA.
XX AC ADB08322;
XX DT 20-NOV-2003 (first entry)
XX DE Alloiooccus otitis antigenic protein SEQ ID NO:2262.
XX KW Alloiooccus otitidis; antigenic protein; immunogenic; immunisation;
XX KW gene therapy; Gram-positive bacterium; infection.
XX OS Alloiooccus otitis.
XX PN WO2003048304-A2.
XX PD 12-JUN-2003.
XX PF 25-NOV-2002; 2002WO-US036123.
XX PR 29-NOV-2001; 2001US-0333777P.
XX PR 18-NOV-2002; 2002US-0426742P.
XX PA (AMHP ) WYETH HOLDINGS CORP.
XX PI Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;
XX DR WPI; 2003-505284/47.
XX DR N-PSDB; ADB08321.
XX DR New Alloiooccus otitidis polynucleotides and polypeptides, useful for
PT treating and diagnosing diseases, drug screening assays and monitoring of
PT effects during drug clinical trials.
XX PS Claim 33; SEQ ID NO 2262; 1019pp; English.
XX CC The present invention describes an isolated polynucleotide (I) of
CC Alloiooccus otitidis genomic DNA, which encodes an antigenic protein.
CC Alloiooccus otitidis is a Gram-positive bacterium. Also described: (1)
CC an isolated polypeptide that is encoded by the polynucleotide (I); (2) an
CC expression vector comprising the novel isolated polynucleotide (I), its
CC complement, degenerate variant or fragment; (3) a genetically engineered
CC host cell, transfected, transformed or infected with the vector of (2);
CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
CC composition comprising the polypeptide, its complement, biological
CC equivalent or fragment, or the polynucleotide that is comprised in the
CC expression vector; (6) a pharmaceutical composition comprising the
CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
CC of the polypeptides of (1), their biological equivalent or fragment; (8)
CC immunising against Alloiooccus otitidis by administering to a host the
CC immunogenic composition; (9) detecting and/or identifying Alloiooccus
CC otitidis in the biological sample; (10) a kit comprising a container
CC containing the novel polynucleotide, its degenerate variant or fragment,
CC or the antibody of (4); and (11) producing a polypeptide by culturing the
CC genetically engineered host cell under conditions suitable to produce the
CC polypeptide from the culture. (I) can be used in gene therapy. The
CC polynucleotides, polypeptides, antibodies and compositions of the present
CC invention can be used for treating and diagnosing diseases, drug
CC screening assays and monitoring of effects during drug clinical trials.
CC The polynucleotides are useful for expressing and detecting Alloiooccus
CC otitidis. The present sequence represents an Alloiooccus otitidis
CC antigen protein from the present invention.
XX SQ Sequence 469 AA;

Query Match 7.5%; Score 150; DB 6; Length 469;
Best Local Similarity 23.7%; Pred. No. 1e-05;
Matches 81; Conservative 44; Mismatches 149; Indels 68; Gaps 14;

QY 10 VLLGGDVMILGRGVQDLPHPGKPOLRERYMRDATGVVLAERVNGRIPLPVDWRWGEA 69
Db 91 ISLVGDVMV-----HGQLEAAV-----DEVSGQ-----YDFDPTFPV 124

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 01:42:44 ; Search time 31.1997 Seconds
(without alignments)
909.197 Million cell updates/sec

Title: US-10-617-038-6

Perfect score: 1991

Sequence: 1 VAGNPDVTVLLGGDVMGLGR.....VTRPDNLLEVPAANLTSKE 380

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/protdata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/protdata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/protdata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/protdata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/protdata/1/iaa/PCtUS_COMB.pep.*
6: /cgn2_6/protdata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	860.5	43.2	358	4	US-09-302-540-15396
2	203	10.2	371	4	US-09-302-540-10866
3	169.5	8.5	594	4	US-09-328-352-5490
4	164	8.2	423	4	US-09-302-540-16421
5	149	7.5	169	4	US-09-302-540-14176
6	125.5	6.3	201	4	US-09-107-433-3237
7	112	5.6	961	4	US-09-328-352-7358
8	108.5	5.4	3546	4	US-08-679-279-13
9	107.5	5.4	4861	4	US-09-919-497-70
10	105	5.3	458	4	US-09-302-540-16600
11	104.5	5.2	160	4	US-09-583-110-3449
12	104.5	5.2	421	4	US-09-252-991A-30742
13	103.5	5.2	888	4	US-09-697-022-4
14	103.5	5.2	888	4	US-09-697-022-5
15	103	5.2	382	4	US-09-252-991A-19700
16	103	5.2	568	4	US-09-252-991A-27039
17	103	5.2	1114	2	US-08-576-626A-31
18	101	5.1	543	4	US-09-351-150A-15
19	101	5.1	927	4	US-09-252-991A-16765
20	100	5.0	389	3	US-08-134-001C-4316
21	100	5.0	524	4	US-09-302-540-13665
22	99.5	5.0	3816	3	US-09-428-517-3
23	99.5	5.0	5588	3	US-09-036-987A-6
24	99.5	5.0	5588	3	US-09-370-700-6
25	99.5	5.0	5588	4	US-09-603-207-6
26	98.5	4.9	585	4	US-09-252-991A-30512
27	98.5	4.9	666	4	US-09-252-991A-32542

28	98	4.9	6396	4	US-09-410-551B-72	Sequence 72, Appl
29	98	4.9	6396	4	US-09-940-316B-72	Sequence 72, Appl
30	97.5	4.9	594	4	US-09-252-991A-16852	Sequence 16852, A
31	97.5	4.9	929	4	US-09-697-022-3	Sequence 3, Appl
32	97.5	4.9	6095	3	US-09-144-085-2	Sequence 2, Appl
33	97	4.9	412	4	US-09-252-991A-28034	Sequence 28034, A
34	97	4.9	671	4	US-09-252-991A-31862	Sequence 31862, A
35	96.5	4.8	446	4	US-09-252-991A-21834	Sequence 21834, A
36	96.5	4.8	577	4	US-09-252-991A-28133	Sequence 28133, A
37	96	4.8	298	4	US-09-252-991A-20712	Sequence 20712, A
38	95.5	4.8	520	4	US-09-257-825B-20	Sequence 20, Appl
39	95.5	4.8	762	4	US-09-252-991A-28078	Sequence 28078, A
40	95	4.8	532	4	US-09-252-991A-22039	Sequence 22039, A
41	95	4.8	719	4	US-09-252-991A-28768	Sequence 28768, A
42	95	4.8	949	4	US-09-902-540-15712	Sequence 15712, A
43	95	4.8	1097	4	US-09-252-991A-22579	Sequence 22579, A
44	95	4.8	1376	4	US-09-252-991A-30344	Sequence 30344, A
45	94	4.7	538	4	US-09-252-991A-22427	Sequence 22427, A

ALIGNMENTS

RESULT 1

US-09-902-540-15396
; Sequence 15396, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15396
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15396

Query Match		43.2%;	Score 860.5;	DB 4;	Length 358;
Best Local Similarity		51.3%;	Pred. No. 1e-81;		
Matches 178;		Conservative 44;	Mismatches 116;	Indels 9;	Gaps 5;
QY	17	MLGRGVDQILPHGPKQLRERYMRDATGYVRLAERVNGRIPLPVDWRMPWGEALAVLENT 76			
Db	1	MTGRGIDQVLPHPPGLHEDYLDKARDYVRLAERNGPIRTDVGFGLMGDLDDEMER 60			
QY	77	ATDVCLINLETTITADGEFADRPVCYRMHPDNVPALTALRPHVCALANNHILDFGQGL 136			
Db	61	APDVRILNLETSVTTSEAWMPDKGIHYRMHPENACLTAAIDCCALANNHVLJDMGYPGL 120			
QY	137	TDTVAALAGAGTOSVGAGADLLAARSAL--VTVGHERRVIVGSAVAESSGVPESWAARR 194			
Db	121	LETLATRLCLGTATAGLGEDQEARPPALIDVTGG---RVTVFSFGDSSSGIPVWAATE 177			
QY	195	DRPGVWLIRDPQORDVA--DDVAQAQVLADKRPGDIAIVSMHWGNGWYATAPGVAFARHL 253			
Db	178	DRPGVDLL--PALSPTAVRQLGERVAARKRTGDIIVVASIHWGNGWYAVPDAQRAFARAL 235			
QY	254	ID-AGIDMVHGSSHHPRPIEYRGKPILYCGDVDDYEGIGGHESFRSELRLYLTVT 312			
Db	236	IDAGVDIITHSSHHPRGIEVHRGHLIILYCGCDFLDYEGISGHTAFRDLTLLYLASI 295			
QY	313	DPASGNLISLQMLPLRVSRMRQASQTDTETWLRNTIERISRRFGIR 359			
Db	296	APATGQLTGLRMLPMQLRQLRPHHASRQDAEWLGGVLDROGRDLGTR 342			


```
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 14176
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-14176

Query Match          7.5%; Score 149; DB 4; Length 169;
Best Local Similarity 32.5%; Pred. No. 1.8e-07;
Matches 39; Conservative 19; Mismatches 52; Indels 10; Gaps 13;

QY 212 DDVAAQVLADKRPDIAIVSMHGSNWGYATAPGDVAFARHLIDAGIDMVHGHSSHHPRP 271
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 7 EDIAAA---RAQADIVLPYFHWGREGTYPYQVRLAHVAIDAGAAGVLAGAHPHVLOA 62
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

QY 272 IEIYRKPIYGGDVDDYEGIGHGHSRSELRLLYLTPTPASCNLLISLOWLPLRVSR 331
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 63 MEDYQRPVYISLGNFV-----FGWNPNRDRKGALWKGHGFP-DGGYLSSELPLRTDR 116
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

RESULT 6
US-09-107-433-3237
; Sequence 3237, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Daneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: 60/051553
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3237:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
```

```
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...201
; SEQUENCE DESCRIPTION: SEQ ID NO: 3237:
US-09-107-433-3237

Query Match          6.3%; Score 125.5; DB 4; Length 201;
Best Local Similarity 22.7%; Pred. No. 7e-05;
Matches 48; Conservative 38; Mismatches 88; Indels 37; Gaps 6;

QY 77 ATDVCLINLETTITADGEPADRKPVCYRMHPDNVPALTALRP---HVCALANNHILDFGY 133
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 1 AADLAIGDFEGTINKDHYLAG-----YLLFNAPVEVMDAIKAGYHVLDLAHNNHILDSQI 55
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

QY 134 QGLTDTVAALAGAGIQSVGA-----GADLLAARRSALVTVGHERRVIVGSVAESSGV 186
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 56 EGVISTADIIERKAGITPIGVYTHEPRDQAPLVIKEVNAI-----KVALLAYSYGFNGI 108
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

QY 187 PESWAARRDRPGWILIRDPAQRDVA---DDVAAQVLADKRPDIAIVSMHGSNWGYAT 242
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 109 EQ-----YIFQEDYNYRLSDLNEDKIKAEVERAEKEADITIIMPQMGVRYRLEP 157
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

QY 243 AFGDVAFARHLIDAGIDMVHGHSSHHPRPIE 273
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 158 TTEQKALYHKMIDMGADIIFGGHPHVPESE 189
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

RESULT 7
US-09-328-352-7358
; Sequence 7358, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7358
; LENGTH: 961
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7358

Query Match          5.6%; Score 112; DB 4; Length 961;
Best Local Similarity 21.2%; Pred. No. 0.021;
Matches 97; Conservative 58; Mismatches 158; Indels 144; Gaps 21;

QY 5 PDVTVLLGGDYMLGRGVDQILPHGPKPQ-----LRERYMRDATGYVRLAERVNGRIP 57
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 276 PNEIAKIIGADV-----PDWPQRSVVLANAHGLHARPATHLVNLTKTLCQDIQ 324
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

QY 58 LPVD-----WRWPWGEALAVLENTATDVCLINLETTITAD---GEFADR 98
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 325 VAVDDSSFSVSAKSLRLLALGCKRGQTLRFIAEPETD-AVEALDQVILAVQQGLGEEVA 383
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

QY 99 KPCVYRMHPDNVPALTALRPHVCALANN-----HILD---FGYGLTDTVA 141
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 384 LPVTSEI-PDQLPAQTAI-SNATTIGNNTGIAASSGLAFPGVHVIVKPYQYERM-----S 183
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 437 ---GSGVKVEKENLDI-----ALHAVKNHIQVIAKADASEIKQIFRAHLEMLDDPDLI 487
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

QY 184 SGVPESWAARRDRPGW-----LIRDPAQ--RDVADDDAAQV---LAD 221
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 488 NGVYQKIYQNLSPAPAAWHEHIEAAAKAQBALPDLRAERATDLRDCGRVLTQCGEVAI 547
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:

QY 222 KRPDIAIVSMHGSNWGYATAPGDVAFARHLIDAGIDMVHGHSSHHPRPIEYRKPI 281
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
Db 548 EEPKEPYILIMH-----DVGPSVARLUNKRVAGILTAIGGASAHSAIVARALGIPAI 600
   :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
```



```

; SEQ ID NO 30742
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30742

Query Match      5.2%; Score 104.5; DB 4; Length 421;
Best Local Similarity 25.0%; Pred. No. 0.035;
Matches 73; Conservative 21; Mismatches 101; Indels 97; Gaps 14;

QY 13 GGDVMLGRGVDQILP-HPGKPOLRERY---MRDATGYVRLAERVNGRIPLPVDWRWPMGE 68
Db 118 CGGTGARGFGREDHPQHDLRPLRLRPRYRPLRNADHPGAAAEPTGTGRAVPARGRGFGD 177
QY 69 -----ALAVLENTATDVCLINLETTITADGEFAFADRPVKVCYRMHPDNPVALTA 115
Db 178 AGGDHCRPGRRPALCVR-----CLLGLR-----ADDSLRLPRVRPDAQAGTQ 221
QY 116 LRPHVCALANNHILDFGYQGLTDTVAALAGAGIOSVGAGADILLAARRSAL--VTVGHER 173
Db 222 PAAH-----PGQGHPLAARRHFFHPQPTGHLR- 249
QY 174 VTGVSVAEASSGVSPESWAARRDRPGVWLIR-DPAORDVADDVAQVLAADKRPDIAVSM 232
Db 250 ---GNLPRVPRAPR---RRHRPAPGVQRQHPAHRALGPRPAALGSSGCRFGDVVLG- 301
QY 233 HMGSNMGYATAPGDVAFARHLIDAGIDMV--HGH---SSHPRPIEIVYRGP 279
Db 302 -----ALQHR-TQRRADHVPLRGLRPGNHLRPRVEIVLAVP 337

RESULT 13
US-09-697-022-4
; Sequence 4, Application US/09697022
; Patent No. 6627427
; GENERAL INFORMATION:
; APPLICANT: Kosan Biosciences, Inc.
; APPLICANT: Katz, Leonard
; APPLICANT: Revill, Peter
; TITLE OF INVENTION: PRODUCTION OF POLYKETIDES
; FILE REFERENCE: 30062-20048.00
; CURRENT APPLICATION NUMBER: US/09/697,022
; CURRENT FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: US 60/161,414
; PRIOR FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Saccharopolyspora erythraea
US-09-697-022-4

Query Match      5.2%; Score 103.5; DB 4; Length 888;
Best Local Similarity 24.0%; Pred. No. 0.14;
Matches 80; Conservative 38; Mismatches 93; Indels 123; Gaps 20;

QY 10 VLLGGDVMLGRGVDQ---ILPHGK-POLR-----ERYMRDATGYVRLAERVNGRIPLPVD 61
Db 504 VLLDDVVVTGTALTEPRTVFVFPGGQPGWRGNGVLMMAASPVAARMQCADALIP-HTG 562
QY 62 WRWPMGEALAVLENTATDVCLINLETTITADGEFAADR---KPCYRMHPDNPVALTA-- 115
Db 563 W-----DPIAMLD-----DPEVTRRVDDVHVPCWAV---MVSLAAVM 596
QY 116 ---LRPHVCALANNHILDFGY-QGLTDTVAALAGAGIOSVGAGADILLAARRSALVTVCH 170
Db 597 EAAGVRPDVAV-----IGHSQG---EIAAACVAGALTLEDGARLVALRSVLLRL 644
QY 171 ERRVIVGSVA-----AESGVPESWAARRD-----RP-----GWL 201
Db 645 AGRGANGSVLPADVEADAARIDGIVWVWAGRNGATTTTVAGRPDAVETLIADYEARGWV 704
```

```

; SEQ ID NO 30742
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30742

Query Match      5.2%; Score 104.5; DB 4; Length 421;
Best Local Similarity 25.0%; Pred. No. 0.035;
Matches 73; Conservative 21; Mismatches 101; Indels 97; Gaps 14;

QY 13 GGDVMLGRGVDQILP-HPGKPOLRERY---MRDATGYVRLAERVNGRIPLPVDWRWPMGE 68
Db 118 CGGTGARGFGREDHPQHDLRPLRLRPRYRPLRNADHPGAAAEPTGTGRAVPARGRGFGD 177
QY 69 -----ALAVLENTATDVCLINLETTITADGEFAFADRPVKVCYRMHPDNPVALTA 115
Db 178 AGGDHCRPGRRPALCVR-----CLLGLR-----ADDSLRLPRVRPDAQAGTQ 221
QY 116 LRPHVCALANNHILDFGYQGLTDTVAALAGAGIOSVGAGADILLAARRSAL--VTVGHER 173
Db 222 PAAH-----PGQGHPLAARRHFFHPQPTGHLR- 249
QY 174 VTGVSVAEASSGVSPESWAARRDRPGVWLIR-DPAORDVADDVAQVLAADKRPDIAVSM 232
Db 250 ---GNLPRVPRAPR---RRHRPAPGVQRQHPAHRALGPRPAALGSSGCRFGDVVLG- 301
QY 233 HMGSNMGYATAPGDVAFARHLIDAGIDMV--HGH---SSHPRPIEIVYRGP 279
Db 302 -----ALQHR-TQRRADHVPLRGLRPGNHLRPRVEIVLAVP 337

RESULT 13
US-09-697-022-4
; Sequence 4, Application US/09697022
; Patent No. 6627427
; GENERAL INFORMATION:
; APPLICANT: Kosan Biosciences, Inc.
; APPLICANT: Katz, Leonard
; APPLICANT: Revill, Peter
; TITLE OF INVENTION: PRODUCTION OF POLYKETIDES
; FILE REFERENCE: 30062-20048.00
; CURRENT APPLICATION NUMBER: US/09/697,022
; CURRENT FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: US 60/161,414
; PRIOR FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 888
; TYPE: PRT
; ORGANISM: Saccharopolyspora erythraea
US-09-697-022-4

Query Match      5.2%; Score 103.5; DB 4; Length 888;
Best Local Similarity 24.0%; Pred. No. 0.14;
Matches 80; Conservative 38; Mismatches 93; Indels 123; Gaps 20;

QY 10 VLLGGDVMLGRGVDQ---ILPHGK-POLR-----ERYMRDATGYVRLAERVNGRIPLPVD 61
Db 504 VLLDDVVVTGTALTEPRTVFVFPGGQPGWRGNGVLMMAASPVAARMQCADALIP-HTG 562
QY 62 WRWPMGEALAVLENTATDVCLINLETTITADGEFAADR---KPCYRMHPDNPVALTA-- 115
Db 563 W-----DPIAMLD-----DPEVTRRVDDVHVPCWAV---MVSLAAVM 596
QY 116 ---LRPHVCALANNHILDFGY-QGLTDTVAALAGAGIOSVGAGADILLAARRSALVTVCH 170
Db 597 EAAGVRPDVAV-----IGHSQG---EIAAACVAGALTLEDGARLVALRSVLLRL 644
QY 171 ERRVIVGSVA-----AESGVPESWAARRD-----RP-----GWL 201
Db 645 AGRGANGSVLPADVEADAARIDGIVWVWAGRNGATTTTVAGRPDAVETLIADYEARGWV 704
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:14:10 ; Search time 90.8729 Seconds
(without alignments)
1742.576 Million cell updates/sec

Title: US-10-617-038-6
Perfect score: 1991
Sequence: 1 VAGNPDVTVLLGGDVMLGR.....VTRPDNLLVVPAAANTSKE 380

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications_AA.*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
 - 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
 - 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1991	100.0	380	15	US-10-617-038-6
2	200	10.0	382	14	US-10-156-761-14822
3	179	9.0	398	18	US-10-501-282-60
4	172	8.6	382	18	US-10-501-282-58
5	153.5	7.7	249	18	US-10-501-282-56
6	150	7.5	469	18	US-10-501-282-2262
7	150	7.5	474	18	US-10-501-282-2264
8	128	6.4	666	10	US-09-746-660A-46
9	128	6.4	666	17	US-10-494-675-120
10	125.5	6.3	201	18	US-10-617-320-3237
11	115.5	5.8	170	17	US-10-472-928-2958

12	115	5.8	2341	17	US-10-844-716-5	Sequence 5, Appli
13	112.5	5.7	430	16	US-10-474-792-54	Sequence 54, Appl
14	112.5	5.7	477	15	US-10-425-114-37484	Sequence 37484, A
15	110	5.5	606	14	US-10-156-761-7747	Sequence 7747, Ap
16	108	5.4	921	14	US-10-156-761-9301	Sequence 9301, Ap
17	107.5	5.4	4861	9	US-09-919-497-70	Sequence 70, Appl
18	107.5	5.4	4861	14	US-10-097-534-26	Sequence 26, Appl
19	107.5	5.4	4861	14	US-10-146-473-49	Sequence 49, Appl
20	107.5	5.4	4861	18	US-10-287-436A-486	Sequence 486, App
21	107.5	5.4	4861	18	US-10-287-436A-1182	Sequence 1182, Ap
22	105	5.3	497	14	US-10-156-761-9418	Sequence 9418, Ap
23	105	5.3	1822	15	US-10-432-443-39	Sequence 39, Appl
24	103.5	5.2	470	11	US-09-855-604-896	Sequence 896, App
25	103.5	5.2	470	12	US-09-855-604-896	Sequence 896, App
26	103.5	5.2	486	11	US-09-855-604-701	Sequence 701, App
27	103.5	5.2	486	12	US-09-855-604-701	Sequence 701, App
28	103.5	5.2	503	14	US-10-156-761-13977	Sequence 13977, A
29	103.5	5.2	738	15	US-10-369-493-20530	Sequence 20530, A
30	103.5	5.2	888	15	US-10-607-809-4	Sequence 4, Appli
31	103.5	5.2	888	15	US-10-607-809-5	Sequence 5, Appli
32	103.5	5.2	1927	14	US-10-152-886-63	Sequence 63, Appl
33	103.5	5.2	1927	20	US-11-053-576-63	Sequence 63, Appl
34	103.5	5.2	1927	20	US-11-053-576-63	Sequence 63, Appl
35	103	5.2	558	14	US-10-032-201B-267	Sequence 267, App
36	102.5	5.1	437	15	US-10-424-599-170690	Sequence 170690,
37	102	5.1	952	15	US-10-282-122A-44862	Sequence 44862, A
38	102	5.1	8438	17	US-10-937-730A-4	Sequence 4, Appli
39	101.5	5.1	534	10	US-09-802-208B-4	Sequence 4, Appli
40	101.5	5.1	704	14	US-10-156-761-12458	Sequence 12458, A
41	101.5	5.1	10625	17	US-10-819-386A-3	Sequence 3, Appli
42	101	5.1	1743	17	US-10-819-386A-2	Sequence 2, Appli
43	100.5	5.0	636	14	US-10-205-032-28	Sequence 28, Appl
44	100.5	5.0	825	15	US-10-369-493-13754	Sequence 13754, A
45	100.5	5.0	833	14	US-10-156-761-14271	Sequence 14271, A

ALIGNMENTS

RESULT 1
US-10-617-038-6
; Sequence 6, Application US/10617038
; Publication NO. US20040057963A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Peter
; APPLICANT: Rosenkrands, Ida
; APPLICANT: Stryhn, Anette
; TITLE OF INVENTION: Therapeutic TB Vaccine
; FILE REFERENCE: SS15AUSA
; CURRENT APPLICATION NUMBER: US/10/617,038
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: DK PA 2002 01098
; PRIOR FILING DATE: 2002-07-13
; PRIOR APPLICATION NUMBER: US 60/401,725
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-617-038-6

Query Match	100.0%	Score	1991:	DB	15:	Length	380;
Best Local Similarity	100.0%	Pred. No.	7.7e-18;				
Matches	380;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	VAGNPDVTVLLGGDVMLGRGVDQILPHFGKQLRERYMRDATGYVRLAERVNGRIPLPV	60				
Db	1	VAGNPDVTVLLGGDVMLGRGVDQILPHFGKQLRERYMRDATGYVRLAERVNGRIPLPV	60				
Qy	61	DWRWPGEALAVLENTATDVCLINLETTTATGCEFAADRPVCVYRMHPDNNVPALTALRPHV	120				

Db 61 DWRPWEALAVLENTATDVCLINLETTITADGEFADRPVCMRHPDNVPALTALRPHV 120
QY 121 CALANNHILDRGYOGLTDTVAALAGAGTQSVGAGADLLAARRSALVTGVGHERRVIVGSA 180
Db 121 CALANNHILDRGYOGLTDTVAALAGAGTQSVGAGADLLAARRSALVTGVGHERRVIVGSA 180
QY 181 AESSGVPSWAARRDRPGVWLIRDPAQRDVAADVAAQVLADKRPGLDIAIVSMHWSNGWY 240
Db 181 AESSGVPSWAARRDRPGVWLIRDPAQRDVAADVAAQVLADKRPGLDIAIVSMHWSNGWY 240
QY 241 ATAPGDVAFARHLIDAGIDMVHGSHHPRPIEIVRGKPIILYGCDDVVDDYEGIGGHESF 300
Db 241 ATAPGDVAFARHLIDAGIDMVHGSHHPRPIEIVRGKPIILYGCDDVVDDYEGIGGHESF 300
QY 301 RSELRLLVLTVDTPASGNLSLQMLPLRVSRMRLQASQDTTEWLNRNTERISRRFGIRV 360
Db 301 RSELRLLVLTVDTPASGNLSLQMLPLRVSRMRLQASQDTTEWLNRNTERISRRFGIRV 360
QY 361 VTRPDNLLEVVPAAULTSKE 380
Db 361 VTRPDNLLEVVPAAULTSKE 380

RESULT 2

US-10-156-761-14822
; Sequence 14822, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14822
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14822

Query Match 10.0%; Score 200; DB 14; Length 382;
Best Local Similarity 24.4%; Pred. No. 2.2e-10;
Matches 99; Conservative 49; Mismatches 153; Indels 104; Gaps 20;
QY 2 AGNPDV-----TVLLGGDVMILGRGVQDQILPHPGPKPOLRERYMRDATGYVRLAERVNVR 55
Db 33 AGGPAAPATRGFTLVASGDV-----JPH---APVIEQAKADAG-----69
QY 56 IPLPVDWRMPWGEALAVLENTAT--DVCLINLETTITADGEFADRPVCMRHPDNVPAL 113
Db 70 --IGYDFR----PMLAAAPVVRADVAICHLETVFEGEDGYSGYS--AFVSPQVALGL 121
QY 114 TALRPHVCALANNHILDRGYOGLTDTVAALAGAGTQSVGAGADLLAARRSALVTGVGHERR 173
Db 122 AAIGYDGGSTASNHALDYDTAGIHSTLNALDRAGVQHTGSAR---LAPEAAAPAMFHADG 178
QY 174 VIVGSA--AESSGVV-----ESWAARR--DRPGVWLIRDPAQRDVAADVAAQVLADKRPGLD 226
Db 179 AQVAHLATHDINGRPLPAGPFWANLIDR-----DRIVADARAARAAG-----AD 224
QY 227 IATVSMHWSNGWYATAPGDVAFARHLIDA-----GIDMVHGSHHPRPIEIVRGKPI 280
Db 225 VVWVSLGWTEWQDEPDQAQLRLSRELTASQTKGRPDIDLILGTRAHVPPQVPEKVGWTW 284

QY 281 LYCGDDVVDDYEGIGGHESFSELRLLVLTVDTPASGNLSLQMLPLRVSRMRLQASQT 340
Db 285 VYGGDQIAG--EWANREGFR-----DRGOE-----STIGRFTTAPPARA 323
QY 341 DTEWLNRNTERISRRFGI-RVV-----TRP--DNLEVV 371
Db 324 GERQVTKAEFIPQRFDLGRVVDLNAALAQAGTAPVRDRIRDVV 368

RESULT 3

US-10-501-282-60
; Sequence 60, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOTOCOCCLUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; PRIOR FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 60
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Alloiococcus otitidis
US-10-501-282-60

Query Match 9.0%; Score 179; DB 18; Length 398;
Best Local Similarity 20.9%; Pred. No. 2.5e-08;
Matches 83; Conservative 61; Mismatches 163; Indels 90; Gaps 11;

QY 1 VAGNPDVVTVLLGGDVMILGRGVQDQILPHPGPKPOLRERYMRDATGYVRLAERVNVRIPLPV 60
Db 65 VQANAETPSISFVGDISLSRVIAIEIGERDGYEFP-----FTYVKPHFDGR-----109
QY 61 DWRPWEALAVLENTATDVCLINLETTITAD-----GEFADRPVCMRHP 107
Db 110 -----DLVVANLESADVTDNRTYVVKPASSLSGGIYLDSS-----1144
QY 108 DNVPAALTALRPHVCALANNHILDRGYOGLTDTVAALAGAGTQSVGAGADLLAARRSALVT 167
Db 145 DSVKAMQEAQIDLVSMANNHTGDMKGQVDMGBILRDVEMDIYIGMGHDAEAGQAYQFV 204
QY 168 VGHERRVIVGSAVVAESSGVPSWAARRDRPGVWLIRDPAQRDVAADVAAQVLADKRPGLD 227
Db 205 ANN-----LTSIFAVSDVIKPGQAASODEPGLVLTNSQAFUNLANSY-----DQESDL 253
QY 228 AIVSMHWSNGWYATAPGDVAFARHLIDAGIDMVHGSHHPRPIEIVRGKPIILYGCDDV 287
Db 254 VIATIHAGIEVIRPDANHQELAESLIDAGADIVICSHTSHLLPVEKYQDGIIFYGLGNF 313
QY 288 VDDYEGIGGHESFSELRLLVLTVDTPASGNLSLQMLPLRV-----SRMRLOAS 338
Db 314 IFDQ----GMOSSDTSV-ILDMDINSP---DQVFTLRPMKIEAGIKPENPNGLVTRIQ 365
QY 339 QTDTEWLNRNTERISRRFGIRVVRPDLNLEVVVPAAN 375
Db 366 KRLTERLDDQDFKID-----RADNIVIDFDLAN 393

RESULT 4

US-10-501-282-58


```

; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; PRIOR FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Alloiococcus otitidis
US-10-501-282-56

Query Match          7.7%; Score 153.5; DB 18; Length 249;
Best Local Similarity 23.6%; Pred.No.3.9e-06;
Matches 62; Conservative 43; Mismatches 121; Indels 37; Gaps 7;

QY 122 ALANNHILDFGYOGLTDTVAALAGAGIQSVGAGADLLAARRSALVTVGHERRVIVGSVAA 181
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 10 SWANNHTGDMGKOGMVDGMEILRDVMDYITGMGHDRAEAGCAQYQFVANN-----LTTSIFA 65

QY 182 ESSGVPE$WAAREDRPGVWLIRDPQAORDVADVAQAQVLADKRPQDIAIVSMHWG$NNGYA 241
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 66 VSDVIRPGQAASDDEPGVLTTSNQAFNLANSY-----DQESDLVIAYIHAGIEYIRQ 118

QY 242 TAPGDVAFARHLIDAGIDMVGHSSHHPRPIEYRGPILYGGCDVVDDYEGIGGHESFR 301
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 119 PDANHQELAESLIDAGADIVIC$H$TH$LL$PVEKYQDGIIFYGLGNFIDQ-----GMQSST 174

QY 302 SEURLLYLVTPDASGNLISLQMLPLRV-----SRMLQRA$OTDTEWLRNTIERI 352
    :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 175 DSV-ILDM$IN$P---DOVRFTLRPMKIEAGIPKPNPGLV$TORIQKRLTERLDDQDFKI 230

QY 353 SRRFGIRVVRPDNLLEVVPAAN 375
    :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 231 D-----RADNIVIDFDLAN 244

RESULT 6
US-10-501-282-2262
; Sequence 2262, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOIOCOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; TITLE OF INVENTION: POLYPEPTIDE ANTIGENS, IMMUNOGENIC COMPOSITIONS AND USES THEREOF
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; PRIOR FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2262
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Alloiococcus otitidis
US-10-501-282-2262

Query Match          7.5%; Score 150; DB 18; Length 469;
Best Local Similarity 23.7%; Pred.No.2e-05;
Matches 81; Conservative 44; Mismatches 149; Indels 68; Gaps 14;

QY 10 VLLGGDVMLGRGVQDILPHPKPOLRERYMRDATGYVRLAERVNGRIPLPVDW$RWPNG$EA 69
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Db 91 ISLVGDVNV-----HGPQLEAAY-----DEVSGQ-----YDFDPTFAPV 124
QY 70 LAVLENTATDVCLINLETTITADGFEADRKPVCYRMHDPNVPALTALRPHVCALANNHIL 129
Db 125 APLIEE--SDLALANLETTLPGE-EFGYTGYPLEFASPSLDALQESGFDGLFTVNNHCL 181
QY 130 DFGVGLTDTVAALAGAGIQSVGAGADLLAAR-----RSALVTVGHERRVIVGSVAE 182
Db 182 DTGVOGLERTAQVDEKGMDAIGTFQDPDQARVLHKDIEGIKVAILLAYTEMVNPQSQTTE 241
QY 183 SSGVPESWAARRDRPGVWLIRDPARDVADVAQVLAD----KRPGLDIAIVSMHWSGNW 238
Db 242 S---PEDWT-----PYNLNLSDE---QILADLDEVEDTDTFLAYVHWGEY 282
QY 239 GYATAPGVAFARHLIDAGIDWVGHSHHRRPPIRYGKP-----ILYCGDVDD--YE 292
Db 283 TDQPTANQVHYAQFLAQGVDAVIGSHPHVIQETEVIEGDOETFPVLYSMGNFASNQVRE 342
QY 293 GIGGHESFRSELRLLYLT-VTDPASG--NLISLQWMLPLRVSR 331
Db 343 ALGQDYAAATEDGVIVSLNLVKDPETGKVDLASVDYDPTWVWR 384

RESULT 7

US-10-501-282-2264
; Sequence 2264, Application US/10501282
; Publication No. US20050203280A1
; GENERAL INFORMATION:
; APPLICANT: MCMICHAEL, JOHN CALHOUN
; APPLICANT: ZAGURSKY, ROBERT JOHN
; APPLICANT: RUSSELL, DAVID PARRISH
; APPLICANT: FLETCHER, LEAH DIANE
; TITLE OF INVENTION: ALLOIOCOCCUS OTITIDIS OPEN READING FRAMES (ORFS) ENCODING
; TITLE OF INVENTION: POLYPEPTIDE ANTIGENS, IMMUNOGENIC COMPOSITIONS AND USES THEREOF
; FILE REFERENCE: AM100780 L2
; CURRENT APPLICATION NUMBER: US/10/501,282
; PRIOR FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: 60/333,777
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 60/426,742
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US02/36123
; PRIOR FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2264
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Alloiococcus otitidis
US-10-501-282-2264

Query Match 7.5%; Score 150; DB 18; Length 474;
Best Local Similarity 23.7%; Pred. No. 2.1e-05;
Matches 81; Conservative 44; Mismatches 149; Indels 68; Gaps 14;
QY 10 VLLGGDVMLGRGVDOILPHPKPOLRERYMRDATGYVRLAERVNGRIPLPVDWRWPWGEA 69
Db 96 ISLVGDVNV-----HGPQLEAAY-----DEVSGQ-----YDFDPTFAPV 129
QY 70 LAVLENTATDVCLINLETTITADGFEADRKPVCYRMHDPNVPALTALRPHVCALANNHIL 129
Db 130 APLIEE--SDLALANLETTLPGE-EFGYTGYPLEFASPSLDALQESGFDGLFTVNNHCL 186
QY 130 DFGVGLTDTVAALAGAGIQSVGAGADLLAAR-----RSALVTVGHERRVIVGSVAE 182
Db 187 DTGVOGLERTAQVDEKGMDAIGTFQDPDQARVLHKDIEGIKVAILLAYTEMVNPQSQTTE 246
QY 183 SSGVPESWAARRDRPGVWLIRDPARDVADVAQVLAD----KRPGLDIAIVSMHWSGNW 238
Db 247 S---PEDWT-----PYNLNLSDE---QILADLDEVEDTDTFLAYVHWGEY 287
QY 239 GYATAPGVAFARHLIDAGIDWVGHSHHRRPPIRYGKP-----ILYCGDVDD--YE 292

Db 288 TDQPTANQVHYAQFLAQGVDAVIGSHPHVLIQETEVIEGDOETFPVLYSMGNFASNQVRE 347
QY 293 GIGGHESFRSELRLLYLT-VTDPASG--NLISLQWMLPLRVSR 331
Db 348 ALGQDYAAATEDGVIVSLNLVKDPETGKVDLASVDYDPTWVWR 389
RESULT 8
US-09-746-660A-46
; Sequence 46, Application US/09746660A
; Publication No. US20030049804A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroeger, Burkhard
; APPLICANT: Schroeder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; APPLICANT: Kim, Jun-Won
; APPLICANT: Lee, Heung-Schick
; APPLICANT: Hwang, Byung-Joon
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
; FILE REFERENCE: BGI-121CP2
; CURRENT APPLICATION NUMBER: US/09/746,660A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/606740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 09/603124
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 46
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-746-660A-46

Query Match 6.4%; Score 128; DB 10; Length 666;
Best Local Similarity 23.4%; Pred. No. 0.0044;
Matches 103; Conservative 53; Mismatches 148; Indels 136; Gaps 25;
QY 10 VLLGGDVMLGRGVDOILPHPKPOLRERYMRDATGYVRLAERVNGRIPLPVDWRWPWGEA 69
Db 195 VVLGNDT--LSRSLHDII-----PKWARRVIRDASTY---PDRVHGTPLPARLE-PWAEK 244
QY 70 LAVLENTATDVCLINLETTITADGFEADRKPVCYRMHDPNVPALTALRPHVCALANNHIL 129
Db 245 LT--SDPATCRHLIE-----EFGSPVNV---LHSGSMPRNI-----NELV 279
QY 130 DFGVGLTDT---VAALAGAG-----IQSVGAGADLLAARRSALVT---YGHERRVIVG 177
Db 280 DAGIQMGVDTRIFPARKANKGLTFVDAVKOTGHGVDVASERELSQVLNRRGVPGERIILSA 339
QY 178 SVAESGVPESWAARRDRPGVWLIRDPARDVAD-----DVAQVLADKRPGDIAIV 230
Db 340 AIK-----PDRLLALAIENGVIISVD--SRDELDRISALVGDRVARVAPRVPDPAVLP 391
QY 231 SMHWG---SNWGYATAGDVAFARHLIDA--GIDMV-----HGHSS----- 266
Db 392 PTRFGERAADWG-----NRLTEVIPGVDIVGLVHLHGYAAKDRALALQECQOL 440
QY 267 -----HHPRPRIEYRGKPILYCGCDVVDDYE-----GIGGHESFR---SE 303

Db 441 VDSLRECGHSPQIDLGGGVPMY-----IESEDMIRYQSAKSATSAGYASFTWKDDP 495

QY 304 LRLLYLTVTPDASGNLISQMLPLRVSRMLQASQTDTEWLNTERISRRFGIR-----359

Db 496 LSNITYPFYQTPVRGNWLK-DVLSKQVAQMLIDRGLRLHLEPGRSLLDGC-----GVTLAEV 550

QY 360 --VVTRPDNLLEVVPAAHLT 377

Db 551 AFVKTRSDGLPLVGLAMNRT 570

RESULT 9

US-10-494-675-120

; Sequence 120, Application US/10494675

; Publication No. US20050019877A1

; GENERAL INFORMATION:

; APPLICANT: Zelder, Oskar

; APPLICANT: Pompejus, Markus

; APPLICANT: Schroder, Hartwig

; APPLICANT: Kroger, Burkhard

; APPLICANT: Klopffrogge, Corinna

; APPLICANT: Haberhauer, Gregor

; TITLE OF INVENTION: Genes coding for metabolic pathway proteins

; FILE REFERENCE: BGI-163US

; CURRENT APPLICATION NUMBER: US/10/494,675

; CURRENT FILING DATE: 2004-05-04

; PRIOR APPLICATION NUMBER: PCT/EP02/12141

; PRIOR FILING DATE: 2002-10-31

; PRIOR APPLICATION NUMBER: DE 101 54 292.1

; PRIOR FILING DATE: 2001-11-05

; NUMBER OF SEQ ID NOS: 164

; SEQ ID NO 120

; LENGTH: 666

; TYPE: PRT

; ORGANISM: Corynebacterium glutamicum

US-10-494-675-120

Query Match 6.4%; Score 128; DB 17; Length 666;

Best Local Similarity 23.4%; Pred. No. 0.0044;

Matches 103; Conservative 53; Mismatches 148; Indels 136; Gaps 25;

QY 10 VLLGGDVMLGRGVQDQLPHPGKPOLRERYMRDATGVVLAERVNGRIPLPVDMRWPMGEA 69

Db 195 VLGNDET-LSRSLHDI-----PKWARRVIRDASTY---PDRVHGTPPLPARLE-PWAEK 244

QY 70 LAVLENTATDVCLINLETTITADGEFADRKPCVYRMHPDNVPALTALRPHVCALANNHIL 129

Db 245 LT--SDPATCRHLE-----EFGSPVNV---LHSGMPNRI-----NELV 279

QY 130 DFGYQGLTDT---VAALAGAG-----IQSVGAGADLLAARRSALVT---VGHERRVIVG 177

Db 280 DAGIQMGVDTRIIPFARKANKGLTFDVAVKDTGHGVDVASERELSQVLNRGVPGERIILSA 339

QY 178 SVAAESGVPESWAARRDRPGWLIRDPQORDVAD-----DVAQVLADKRPQDIAIV 230

Db 340 AIK-----PDRLLALAIENGVIISVD--SRDELRIISALVGDRVARVAPRVPDPAVL 391

QY 231 SMHWG---SNWGYATAPGDVAFARHLIDA--GIDMV-----HGHSS-----266

Db 392 PTRFGERAADWG-----NRLTEVIPGVDIVGLVHLHLGYAAKDBALALQECQOL 440

QY 267 -----HHPRIETIRGKPILYCGDGVDDYE-----GIGHSESFR---SE 303

Db 441 VDSLRECGHSPQIDLGGGVPMY-----IESEDMIRYQSAKSATSAGYASFTWKDDP 495

QY 304 LRLLYLTVTPDASGNLISQMLPLRVSRMLQASQTDTEWLNTERISRRFGIR-----359

Db 496 LSNITYPFYQTPVRGNWLK-DVLSKQVAQMLIDRGLRLHLEPGRSLLDGC-----GVTLAEV 550

QY 360 --VVTRPDNLLEVVPAAHLT 377

Db 551 AFVKTRSDGLPLVGLAMNRT 570

RESULT 10

US-10-617-320-3237

; Sequence 3237, Application US/10617320

; Publication No. US20050136404A1

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID

; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE

; THERAPEUTICS

; NUMBER OF SEQUENCES: 5206

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM ISO9660

; COMPUTER: <Unknown>

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: <Unknown>

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/617,320

; FILING DATE: 10-Jul-2003

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,433

; FILING DATE: 30-Jun-1998

; APPLICATION NUMBER: 60/ 085131

; FILING DATE: May 12, 1998

; APPLICATION NUMBER: 60/051553

; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Ariniello, Pamela Deneke

; REGISTRATION NUMBER: 40,489

; REFERENCE/DOCKET NUMBER: GTC-011

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (781)893-5007

; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 3237:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 201 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: YES

; ORIGINAL SOURCE:

; ORGANISM: Streptococcus pneumoniae

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (B) LOCATION 1...201

; SEQUENCE DESCRIPTION: SEQ ID NO: 3237:

US-10-617-320-3237

Query Match 6.3%; Score 125.5; DB 18; Length 201;

Best Local Similarity 22.7%; Pred. No. 0.0015;

Matches 48; Conservative 38; Mismatches 88; Indels 37; Gaps 6;

QY 77 ATDVCLINLETTITADGEFADRKPCVYRMHPDNVPALTALRP---HVCALANNHILDG 133

Db 1 AADLAIGDFEGTINKDHYLAG-----YLLFNAPVEVMDAIKEAGYHVLDAHNNHILDSQ 55

QY 134 QGLTDTVAALAGAGTSVGA-----GADLLAARRSALVTVGHERRVIVGSVAAESGV 186

Db 56 EGVIETADIIKAGITPIGVYTHEPRDQAPLVIKENVAI-----KVALLAYSVCNGI 108

QY 187 PESWAARRDRPGWLIRDPQORDVA---DDVAAQVLADKRPQDIAIVSMHWGSMGYAT 242

Db 109 EQ-----YIFQEDYNRYLSDLNEDKTKAEVERAEKEADITIIMPQMGVEYRLEP 157

QY 243 AFGDVAFARHLIDAGIDMVHGHSSHHPRPIE 273

Db 158 TEEQKALYHKMIDWAGADIIFGQHPHVPESE 188

RESULT 11

US-10-472-928-2958

; Sequence 2958, Application US/10472928

; Publication No. US20050020813A1

; GENERAL INFORMATION:

; APPLICANT: CHIRON SpA

; APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH

; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS

; FILE REFERENCE: P026926W0

; CURRENT APPLICATION NUMBER: US/10/472,928

; CURRENT FILING DATE: 2003-09-26

; PRIOR APPLICATION NUMBER: GB-0107658.7

; PRIOR FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 4979

; SOFTWARE: SeqWin99, version 1.03

; SEQ ID NO 2958

; LENGTH: 170

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae

; FEATURE:

; OTHER INFORMATION: capA-related protein

; OTHER INFORMATION: Cellular location: cytoplasm

; OTHER INFORMATION: Similar to strain R6 sequence 15903351 (2.E-95)

US-10-472-928-2958

Query Match 5.8%; Score 115.5; DB 17; Length 170;

Best Local Similarity 23.1%; Pred. No. 0.011;

Matches 39; Conservative 25; Mismatches 70; Indels 35; Gaps 4;

QY 119 HVCALANNHILDFGQGLTDTVAALAGAGIQSVGAGADLLAARRSALVTVCHERRVIVGS 178

DB 10 HVLDAHNNHILDSQIEGVISTADIIEKAGITP-----VYTHEPRDQAPL 55

QY 179 VAAESSGVP-----ESWAARRDRGVWLIRDPADRDVADD-VAAQVLADKRP 224

DB 56 VKEVNGIKVALLAYSQFNGIEQISQED-----YNYLSLNDNEKMKVETERAEK 108

QY 225 GDIAIVSMHWSNGYATAPGDVAFARHLIDAGIDMVHGHSHHPRPIE 273

DB 109 ADITIIMLQMGVEYRLPTEEQKALYHKMIDWAGADIIFGQHPHVPESE 157

RESULT 12

US-10-844-716-5

; Sequence 5, Application US/10844716

; Publication No. US20050003409A1

; GENERAL INFORMATION:

; APPLICANT: Huang, Chengjin

; APPLICANT: Chaleff, Deborah T.

; APPLICANT: Ruppen, Mark E.

; APPLICANT: Stephens, Jerome

; TITLE OF INVENTION: Cloning Genes From Streptomyces Cyaneogriseus Subsp.

; TITLE OF INVENTION: Noncyanogenus for Biosynthesis of Antibiotics and Methods of Use

; FILE REFERENCE: AM100484

; CURRENT APPLICATION NUMBER: US/10/844,716

; CURRENT FILING DATE: 2004-05-13

; NUMBER OF SEQ ID NOS: 42

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 5

; LENGTH: 2341

; TYPE: PRT

; ORGANISM: bacteria

US-10-844-716-5

Query Match 5.8%; Score 115; DB 17; Length 2341;

Best Local Similarity 24.1%; Pred. No. 0.45;

Matches 96; Conservative 48; Mismatches 161; Indels 94; Gaps 20;

QY 20 RGVDQILPHPGKPOLRERYMRDATGYVRLAERV---NGRIPLPVDWRPWPGEALAVLENT 76

Db 1464 RALTRLLPGRGVPTPRAPWOPHATGLLGPADRAPCGSGLEBHDLCGAWPPPGCAVPLVPGE 1523

QY 77 ATDV--CLINLE-----TTITADGEFADRKPCVCYRHMHPDNVPALT 114

Db 1524 LGDVPGCYARLADEGFYGPAPRGLRAWRRTGTIFAVALPAGDGSVFRHLHPALLDAV- 1592

QY 115 ALRPHVCALANN---HILDFGYQGLTDTVAALAGAGIQSVGAGADLLAARRSALVTVCHE 171

Db 1593 -LHPVVLGLVDGVPARPLPFSWNGV--ALHAPASGALRVRLAPADDDGAVGITAATAAG-E 1638

QY 172 RRVIVGSVAEASSGVPSWAARRDRPGV-----WLIRDPADRDVAD---DVAA----- 216

Db 1639 PVLVAALALRSASAEOLRAAIRSAGSRDALYELDWLPL-PADRAASPGGADI AALGTS 1697

QY 217 -----QVLADKRPGDIAIVSMHWSNGYATAPGDVAFARHLIDAGIDMVH 262

Db 1698 ELPCTYETIAELSQLADGAPADPAV-----SDVGAVGGLPDTVSLHGLCRGLELVQ 1752

QY 263 GHSSHHPRIE-----IVRGKPILYGCG---DVVDVYEGI--GGHESPRSEL--RLLYLT 310

Db 1753 AWLG-EPTADTRVLVTRG---AVGCAPAPPVADPAAALWGLVRSQAQAEHPGRLLLLD 1808

QY 311 V-----TDPASGNLISLQW-----LPLRVSRMLQRASQ 339

Db 1809 LDPAGSRPVSGLVEQAVACGEPIAVRGDLGRVPRLSR 1847

RESULT 13

US-10-474-792-54

; Sequence 54, Application US/10474792

; Publication No. US20040236072A1

; GENERAL INFORMATION:

; APPLICANT: Olmsted, Stephen

; APPLICANT: Zagursky, Robert

; APPLICANT: Nickbarg, Elliot

; APPLICANT: Winter, Lourie

; TITLE OF INVENTION: SURFACE PROTEINS OF STREPTOCOCCUS PYOGENES

; FILE REFERENCE: AM 100399

; CURRENT APPLICATION NUMBER: US/10/474,792

; CURRENT FILING DATE: 2003-10-14

; NUMBER OF SEQ ID NOS: 674

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 54

; LENGTH: 430

; TYPE: PRT

; ORGANISM: Streptococcus pyogenes

US-10-474-792-54

Query Match 5.7%; Score 112.5; DB 16; Length 430;

Best Local Similarity 22.1%; Pred. No. 0.077;

Matches 67; Conservative 51; Mismatches 124; Indels 61; Gaps 14;

QY 2 AGNPDVVT--VLLGGDVMLGRGVDOILPHGPKPOLRERYMRDATGYVRLAERVNGRIPLP 59

Db 76 SAQPKVKTARVVANGDIL---IHDIL-----YMSARKADDTYDFTPYFEYVK----- 119

QY 60 VDRWHPWGEALAVLENTATVCLINLETTITADGEFADRKPCVCYRMH---PDNVP-ALTAL 116

Db 120 -DW-----ISGADLAIGDYGTISPDYPLAG-----YPLFNAPBEIAGALKNT 161

QY 117 RPHVCALANNHILDFGYQGLTDTVAALAGAGIQSVGAGADLLAARRSALVTVCHERRVIV 176

Db 162 GYDVVDLAHNNHILDSQLDGLANTKKVPHOLGIDIGI-YDKDRSKESFLIKNVNGIKITAI 220

QY 177 GSVAEASSGVPSWAARRDRPGVWLIRDPADRDVADDVAAQV-----LADKRPGDIAIVS 231

Db 221 LGSYGYNGMEAT-----LSQEDYEKHSMDLDEAKIKKELQLAEKK-ADVTITVM 268

QY 232 MHWSNGYATAPGDVAFARHLIDAGIDMVHGHSHHPRPIE-IYRG---KPILYGCGDV 287

Db 269 PQMTGEVALEPTAQKELYHKMIDWAGDVVLGGHPHPIETSETVIKGRQKFIITYSMGNF 328

QY 288 VDD 290

Db 329 ISN 331

RESULT 14

US-10-425-114-37484

; Sequence 37484, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E.

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 37484

; LENGTH: 477

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: L1B3075-003-G10_FLI.pep

US-10-425-114-37484

Query Match 5.7%; Score 112.5; DB 15; Length 477;

Best Local Similarity 25.1%; Pred. No. 0.089;

Matches 112; Conservative 34; Mismatches 144; Indels 157; Gaps 26;

QY 5 PDVTVLLGGVNLGRGVQDILPHP-----GKPK-----LRERYMRDATGYVRL 48

Db 21 PDAVGDDLGAVLVAVGVVVCVPLDQGVABAGLPEGRVAFRLRLRE-VWRDALDVAI 79

QY 49 AER-VNGRIPLPVDWNPW-----GEALAVLENTATDVCLINLETTITADGEFADRK 99

Db 80 AYRDVDRGVPCP-DGRAPGVDDIVGDVGHVAAAL-----VRVDVAGDDEVG--- 124

QY 100 PVCYRMHPDNVPALTALRHVCAL-----ANNHILDFGYOGL- 136

Db 125 -----LGGDD-PAL-ELHPLRLHALHVVAVAAPVGRVHEHDQPRGGAVHARQLGPPLP 177

QY 137 TDTVAALAGAGIQS---VGAGADLLAARSALVTUGHERRVIVGSVAABSSGVPESSWAAR 193

Db 178 LGEVLAVGGVGGHDDVRGDDPHRVERRRVAVREGLEHVL-----AQLRGQVQLVVAR 232

QY 194 RDRPGVWLIRDP-----QR-----DVADDVAA-----QVLADKR 223

Db 233 RPHPR-LVRRALHEARPRVPVGVVQVVVGVGVADVQHGVAARLVLRQRRVGVKR 290

QY 224 PGDIA-----IVSMHWSNW-GYATAPGDVAFARHLIDAGIDMVHGHSSHHPRIEI 274

Db 291 LAQVANEAELEGHVRAH---GWRGAEAFHRAEPHGVVVVLG----- 330

QY 275 YRCKPILYG-GDVDDYEG-----IGGHESFSELRL-----YLTVTDPASGNLSLQM 324

Db 331 -----AMLEVCAGVDEAGGVVVVAVRGALRPELHLLVRRLAVHDGAAQLIGRHP 386

QY 325 LPLRVSRMLQRASQDTEWLRNTIER 351

Db 387 ---RHDHRAALLVTQRDVHLLRGIVPR 410

RESULT 15

US-10-156-761-7747

; Sequence 7747, Application US/10156761

; Publication No. US20030119018A1

; GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI

; APPLICANT: IKEDA, HARUO

; APPLICANT: ISHIKAWA, JUN

; APPLICANT: HORIKAWA, HIROSHI

; APPLICANT: SHIBA, TADAYOSHI

; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262

; CURRENT APPLICATION NUMBER: US/10/156,761

; CURRENT FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: JP 2001-204089

; PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697

; PRIOR FILING DATE: 2001-08-02

; NUMBER OF SEQ ID NOS: 15109

; SEQ ID NO 7747

; LENGTH: 606

; TYPE: PRT

; ORGANISM: Streptomyces avermitilis

US-10-156-761-7747

Query Match 5.5%; Score 110; DB 14; Length 606;

Best Local Similarity 22.4%; Pred. No. 0.22;

Matches 105; Conservative 46; Mismatches 158; Indels 160; Gaps 27;

QY 3 GNPVV-----TVL-LGGDVMLGRGVQDILPHFGKPOLRERYMRDATGYV--- 46

Db 178 GNPNTVTRARTIARSGTLVDWQGDV---HAATEVVAIRRTQPATLLTEEAKGTVMWDR 234

QY 47 RLAEVNGRIPLPVD-W-----RWPW-----GEAL-----AVLENTAT 78

Db 235 TRAMRILAGEVVRPESRWHEPVTVPHEFVPRW-WOSDGDWLLRDCGPAPADWLEAAYT 293

QY 79 DVCLINLETTITAD-----GEFADRKEVCYRMHPDNVPALTA-LRPHVCALANNHIL---D 130

Db 294 DQTLVTRVDACPADHAAPCAMASGTPT---SSSTLPSLVVRMYRHAMIAENSVDLVTTG 349

QY 131 FGQGLTDTVAALAGAGIOSVGAGADLLAARSALVTV-GHERRVIVGSVAABSSG--- 185

Db 350 TGY-GTALACARLGHARVTSIDVDADLVKAASDRLVLVAGYRPMQAVGDTGELPGAYDR 408

QY 186 -----VPESWAARRDRPGVWLIRDPARDVDVADVAQAOLADKRPGDIAIVSMHWS 236

Db 409 IATVSVRVPVSVLSAL-RPGRLVTTIAGTGLI-----LAADK-----T 448

QY 237 NMGYATAPGDVAFARHLIDAGIDMVHGHSSHHPRIEYRGKPILYCCGDVDDYEGIGG 296

Db 449 NEGGAT--GRIEW-----DRAGFMPTRHGADYCPAD-----DVMQEAEEADG 489

QY 297 HBSFSELRLLYLTVTDPASGNLSLQMLPLRVSRMLQRAS-----QDTEWLR 346

Db 490 EDVITSRYPLLY-----PPDAWDVMS--MLELQIPGITYRRSMDGGNRTVWLLHPDGSWAR 543

QY 347 NT-----IERISPR-----FGIRVVTRPD 365

Db 544 ATARGELDSPTVHQGFRLWSDLRIRNLNREGALPVYGAQVTTTPD 592

Search completed: October 15, 2005, 04:14:57

Job time : 91.8729 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:59:30 ; Search time 21.2037 Seconds
(without alignments)
1724.342 Million cell updates/sec

Title: US-10-617-038-6
Perfect score: 1991
Sequence: 1 VAGNPDVTVLLGGDVMLGR.....VTRPDNLLVWPAANLTSKE 380
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1988	99.8	380	2 E70933	hypothetical prote
2	252	12.7	380	2 H70069	poly-gamma-glutami
3	234	11.8	374	2 H87450	conserved hypothet
4	232.5	11.7	411	2 C30091	capA protein - Bac
5	198.5	10.0	411	2 E95918	conserved hypothet
6	183	9.2	372	2 AB1077	Bacillus anthracis
7	177	8.9	476	2 AE1139	Bacillus anthracis
8	166	8.3	475	2 AD1497	Bacillus anthracis
9	133.5	6.7	304	2 B75595	conserved hypothet
10	116.5	5.9	7576	2 T17428	PK506 polyketide s
11	116	5.8	211	2 B81117	hypothetical prote
12	115.5	5.8	170	2 B95169	conserved domain p
13	115.5	5.8	170	2 C98035	hypothetical prote
14	113.5	5.7	211	2 A81905	hypothetical prote
15	108.5	5.4	409	1 G69000	molybdenum cofacto
16	107.5	5.4	512	2 F83761	fumarate hydratase
17	107.5	5.4	4861	2 S71752	giant protein p619
18	105	5.3	502	2 T35910	probable carboxyle
19	103.5	5.2	470	2 C70672	hypothetical prote
20	101.5	5.1	648	2 C83023	probable oxidoredu
21	101.5	5.1	919	2 T37062	probable transcrip
22	101	5.1	356	2 F95381	conserved hypothet
23	100.5	5.0	280	2 A99805	hypothetical prote
24	100.5	5.0	280	2 F85613	hypothetical prote
25	100	5.0	494	2 S04406	methoxyneurosporen
26	99.5	5.0	359	2 C70735	probable idSA prot
27	99.5	5.0	397	2 G84247	atrazine chlorohyd
28	99.5	5.0	427	2 D75228	probable L-asparag
29	99.5	5.0	1171	2 T35548	hypothetical prote

ALIGNMENTS

RESULT 1

E70933

hypothetical protein Rv0574c - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: E70933

R: Cole, S. T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M. A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A: Authors: Squares, R.; Sulston, J. E.; Taylor, K.; Whitehead, S.; Barrell, B. G.

A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A: Reference number: A70500; MUID: 98295987; PMID: 9634230

A: Accession: E70933

A: Status: preliminary; nucleic acid sequence not shown; translation not shown

A: Molecule type: DNA

A: Residues: 1-380 <COL>

A: Cross-references: UNIPROT: O53771; GB: AL021942; GB: AL123456; NID: g3242298; PIDN: CAA17445

C: Experimental source: strain H37RV

C: Genetics:

A: Gene: Rv0574c

Query Match 99.8%; Score 1988; DB 2; Length 380;

Best Local Similarity 99.7%; Pred. No. 6.3e-147;

Matches 379; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY	1	VAGNPDVTVLLGGDVMLGRGVDDQILPHPGKPOLRERYMRDATGYVRLAERVNNGRIPLPV	60
DB	1	MAGNPDVTVLLGGDVMLGRGVDDQILPHPGKPOLRERYMRDATGYVRLAERVNNGRIPLPV	60
QY	61	DWRWPGEALAVLENTATDVCLINLETTITADGEPADRPVCYRMHPDNPALTALRPHV	120
DB	61	DWRWPGEALAVLENTATDVCLINLETTITADGEPADRPVCYRMHPDNPALTALRPHV	120
QY	121	CALANNHILDFQYQGLTDTVAALAGAGIQSVGAGADLLAARRSALVTVGHERRVIVGSVA	180
DB	121	CALANNHILDFQYQGLTDTVAALAGAGIQSVGAGADLLAARRSALVTVGHERRVIVGSVA	180
QY	181	ASSSGVPESWAARRRPGVWLIIRDPAQRDVAADVAQVLADKRPGDIAIVSMHWGNGWY	240
DB	181	ASSSGVPESWAARRRPGVWLIIRDPAQRDVAADVAQVLADKRPGDIAIVSMHWGNGWY	240
QY	241	ATAPGDVAFARHLIDAGIDMVHGHSHHPRPIETIYRGKPILYCGGVDDYEGIGGHESF	300
DB	241	ATAPGDVAFARHLIDAGIDMVHGHSHHPRPIETIYRGKPILYCGGVDDYEGIGGHESF	300
QY	301	RSELRLLYTVTDPASGNLISLQMLPLRVSRMLQRASQDTDEWLNTIERISRRFGIRV	360
DB	301	RSELRLLYTVTDPASGNLISLQMLPLRVSRMLQRASQDTDEWLNTIERISRRFGIRV	360
QY	361	VTRPDNLLVWPAANLTSKE 380	

Db 361 VTRPDNLLEVVPAANLTSKE 380

RESULT 2

H70069

poly-gamma-glutamic synthesis PgsA protein - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C:Accession: H70069; JC7099

C:Title: Complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: H70069

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-380 <KUN>

A:Cross-references: UNIPROT:P96738; GB:Z99122; GB:AL009126; NID:G2636029; PIDN:CAB15605.

A:Experimental source: strain 168

R:Ashiuchi, M.; Soda, K.; Misono, H.

Biochem. Biophys. Res. Commun. 263, 6-12, 1999

A:Title: A poly-gamma-glutamate synthetic system of Bacillus subtilis IFO 3336: Gene cld

A:Reference number: JC7097; MUID:99417512; PMID:10486244

A:Accession: JC7099

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-380 <ASH>

A:Cross-references: DBJ:AB016245; NID:G6045071; PIDN:BAAB5265.1; PID:G6045074

A:Experimental source: strain IFO 3336

C:Comment: This protein is a membrane-associated protein, it is essential for the polymere

C:Genetics:

A:Gene: ywtB; pgsA

C:Keywords: membrane-associated protein

Query Match 12.7%; Score 252; DB 2; Length 380;

Best Local Similarity 26.0%; Pred. No. 5.1e-12;

Matches 77; Conservative 51; Mismatches 104; Indels 64; Gaps 11;

QY 6 DVTVLLGGDVMGLRGVDOILPHPKPOLRERYMRDATGYV-----RLAERYNGRIPLPVD 61

DB 57 DVLSASFVDIMMGRYVSKVTEQKQADSIFQ-----YVEPIFRASDYVAGNFENPVT 108

QY 62 WRWPMPGEALVALENTATDVCLINLETTITADGEFADRPKPCVCRMHPDNVPALTLRPHVC 121

DB 109 YQKNYKQA-----DKELHQTNKESVKVLKMDNFTVL 140

QY 122 ALANNHILDFGYQGLTDTVAALAGAGIOSVGAGADLLAARRSALVTVGHERRVIVGSVAA 181

DB 141 NSANNHAMDYGVQGMKDTLGEFAKQNLDIVGAGYSLSDAKKK-----ISYQK---VNGVTI 193

QY 182 ESSGVPE---SWAARRDRPGVWLIRDPDQAORDVADDDVAQVLAD-KRPGDTAIYSMHWS 236

DB 194 ATLGTDTVSGKGFAAKNTPGV-LPADP-----EIFPMISEAKHADIVVQSHWGQ 245

QY 237 NMGYATAPGD--VAFARHLIDAGIDMVHGHSHHPRIETIYRGKPILYCGGDVDD 290

DB 246 E--YDNDPNDRQRLARAMSDAGADIIVGHGHPHVLPIEVYNGTIVFVSLGNFVFD 299

RESULT 3

H87450

conserved hypothetical protein CC1625 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C:Accession: H87450

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.I. B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, n. J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: H87450

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-374 <STO>

A:Cross-references: UNIPROT:Q9A7U4; GB:AE005673; NID:gl3423028; PIDN:AAK23604.1; GSPDB:GH

C:Genetics:

A:Gene: CC1625

Query Match 11.8%; Score 234; DB 2; Length 374;

Best Local Similarity 33.9%; Pred. No. 1.3e-10;

Matches 82; Conservative 27; Mismatches 113; Indels 20; Gaps 8;

QY 59 PVDWRMPGEALVALENTATDVCLINLETTITADGEFADRPK--VCYRMHPDNVPALTLAL 116

DB 43 PKDAQWPGRRAVARALARA-DLGTNLLETVIKGPAGAPTREALTLHTGEPAVLDALKAM 101

QY 117 RPHVCALANNHILDFGYQGLTDTVAALAGAGIOSVGAGADL--LAARRSALVTVGHERRV 174

DB 102 DLRLLATANNHAFDLGGGILDTVDATERAGLVSGSGRDLARAAAPAYATPPAGP---- 157

QY 175 IVGSVAAESGVPSSWAARRDRPGVWLIRDPDPA---QRDVADDDVAQVLADKRPGDIAIV 230

DB 158 -VALVAFATGKVRGGGAATWAPGCVNELRRDASGLPREDDAERVLAAAEAKRAEVVIA 216

QY 231 SMH---WGSNWGYATAPG-DVAFARHLIDAGIDMVHGHSHHPRIETIYRKPILYCGGD 286

DB 217 YQNHNDWEPN--QADVPAWQRAFARRCVAAAGASVFVGHGAPLLOGIEIHGAPLFFGLGN 274

QY 287 VV 288

DB 275 FV 276

RESULT 4

C30091

capA protein - Bacillus anthracis

C:Species: Bacillus anthracis

C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004

C:Accession: C30091

R:Makino, S.I.; Uchida, I.; Takado, N.; Sasakawa, C.; Yoshikawa, M. J. Bacteriol. 171, 722-730, 1989

A:Title: Molecular characterization and protein analysis of the cap region, which is esse

A:Reference number: A91902; MUID:89123146; PMID:2536679

A:Accession: C30091

A:Molecule type: DNA

A:Residues: 1-411 <MAK>

A:Cross-references: UNIPROT:P19579; GB:M24150; NID:gl42630; PIDN:AAA22288.1; PID:gl42633

C:Keywords: transmembrane protein

Query Match 11.7%; Score 232.5; DB 2; Length 411;

Best Local Similarity 25.6%; Pred. No. 1.8e-10;

Matches 75; Conservative 48; Mismatches 113; Indels 57; Gaps 9;

QY 8 VTVLLGGDVMGLRGVDOILPHPKPOLRERYMRDATGYVRLAERVNGRIPLPVDWRWPG 67

DB 60 LTMTMVGDIMMGRHVKEIVNRYG----TDYVFRHVSPYLKNSDYVSGNFEPV----- 108

QY 68 EALVALENTAT-----DVCLINLETTITADGEFADRPKPCVCRMHPDNVPALTLRPHVC 121

DB 109 ----LLEDKKNYQKADKNIHLSAKEETVKVKE-----AGFTVLN----- 144

QY 122 ALANNHILDFGYQGLTDTVAALAGAGIOSVGAGADLLAARRSALVTVGHERRVIVGSVAA 181


```
Db 145 -LANNHMTDYGAKTKDTIKAKKEADLDYVGAGENFKVDKNIVYQNVGRVATLGFDTA 203
QY 182 ESGVPEWAARDRPGVWLIRDPQROVADVAAQVLADKRP-----GDIAIVSMHWSGN 237
Db 204 FVAGA-----IATKEQPGS-LSMNP-----DVLKQISKAKDPKKGNADLVVVTNHWGEE 252
QY 238 WGYATAGDVAFARHLIDAGIMVHCHSSHPRPIEYRGKPILYCCGDVDD 290
Db 253 YDNKPSPROEALAKAMVDAGADIIVGHHPHVLQSFVDYKQGIIFYSLGNFVFD 305

RESULT 5
E95918
Conserved hypothetical protein SMb21041, probable cell wall amino acid processing [imported]
C:Species: Sinorhizobium meliloti
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: E95918
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A>Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95918
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-411 <KUR>
A:Cross-references: UNIPROT:Q92VT6; GB:AL591985; PIDN:CAC49013.1; PID:g15140498; GSPDB:C
A:Experimental source: strain 1021, megaplasmid pSymb
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, L.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A>Title: The complete genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMb21041
A:Genome: plasmid

Query Match 10.0%; Score 198.5; DB 2; Length 411;
Best Local Similarity 22.7%; Pred. No. 8.1e-08;
Matches 98; Conservative 52; Mismatches 146; Indels 135; Gaps 19;

QY 34 LRERYMRDATGYVRLAERNRIPLVDRWRPWGEALAVLENTATDVCLINLETTITA-- 91
Db 1 MNDRTFLAVTGQSLIKHDIRD-IPAPA-----FRELQSLRQ--ADLSFTNFEGLGRH 52
QY 92 -----DGEFADRPVRCYRMHPDNVPALTALRPHVCALANNHILDGFGYQGLTVAALAGA 146
Db 53 GGWPLKGSFPG---C--SDPVVLDTLRAIGFRALSLSNHAFDLGPGSVLSTLEEVEKR 106
QY 147 GIOSVGAGADLLAARSRALVTGHERRVIVGSVAESSGVPE-SWAARED-----RPGVW 200
Db 107 GFLHAGLRDQAVSRASTATIGRR---VAIVAMDGPGPDPMFAADGDDDRPARPGVN 163
QY 201 LIR-----DPA----- 206
Db 164 RLRLTQVLEVDGAFEQIRAIKDGYTAIDLNTDSQDDPRHPADPHEIGIARAVFRS 223
QY 207 -----QRVDADVAAQVLADKRPQGDIAIVSM---HWGSNWGYATAPGDVAFARHL 253
Db 224 GQGRSVKIDEADLSRLNSAVATAAKN-GSLVVAVLHHHWSAW-QVFPDVMVSGVAREC 281
QY 254 IDAGIMVHGHSSHPRPIEYRGKPILYCCGDV-----DDYEGIGGHESF 300
Db 282 IDGAAMFVSHGAPVLOPIEYRGRIIFYSLGNFIHVRSKSTWTAPVWESVGVGCSF 341
QY 301 RSELRLLYITVDPASG--NLISLQMLPLRV-----SRMR-LQASQDTWEIWRNTI 349
Db 342 REDNRLIETLTVVVGDEALADRLRLRLAPHLATGESAARILQRCSEQ----- 392
QY 350 ERISRRFGIRV 360
```

```
Db 393 ---SARLGVDI 400

RESULT 6
AB1077
Bacillus anthracis capA protein (polyglutamate capsule biosynthesis) homolog lmo0017 [im
C:Species: Listeria monocytogenes
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AB1077
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.;
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat
ok, C.; Schlueter, T.; Smoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A>Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1077
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-372 <GLA>
A:Cross-references: UNIPROT:O8YAU6; GB:NC_003210; PIDN:CAC98232.1; PID:g16409376; GSPDB:C
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo0017

Query Match 9.2%; Score 183; DB 2; Length 372;
Best Local Similarity 23.2%; Pred. No. 1.1e-06;
Matches 73; Conservative 51; Mismatches 146; Indels 44; Gaps 11;

QY 29 PGKPOLRERYMRDATGYVRLAERNRIPLVDRW-----RWPWGEALAVLENTA 77
Db 52 PPKP--KEPFNIDFTGDMF-----DWDLRPVLAEKMGDYPFNVRRELK--S 95
QY 78 TDVCLINLETTITADGEFADRPVRCYRMHPDNVPALTALRPHVCALANNHILDGFGYQGLT 137
Db 96 SDYTFVDLETAITTRTKVPYQEFWIKSDPSLTALKNAGVDMVNISNHILDYEDGGL 155
QY 138 DTVAALAGAGIOSVGAGADLLAARSRALVTGHERRVIVGSVAESSGVPESWAARRDRP 197
Db 156 DTTAALRANLAYVGAGKNEDEAYQLKVADIKGNK---VGFMSFCHFFPNTGWIADDP 212
QY 198 GWLIRDPQROVADVAAQVLADKRPQGDIAIVSMHWSNWGYATAPGDVAFARHLIDAG 257
Db 213 GV---TNGYDLNLVBEKIKEERAKNKIDYMWVYFHWGVEKNTNPVDYQTYVVKLVDDN 269
QY 258 -IDMVHGHSSHPRPIEYRGKPILYCGGD-VVDDYEGIGGHESFRSELRLLY-----LT 310
Db 270 LVDAIVASHPHWLQGFYKDVPIAISLGNLFDPDY--VSGHSAETGIYKLNFDQGVTA 327
QY 311 VTDPAA--SGNLISL 322
Db 328 HFDPGIISGNQINM 341

RESULT 7
AE1139
Bacillus anthracis encapsulation protein CapA homolog lmo0516 [imported] - Listeria mono
C:Species: Listeria monocytogenes
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AE1139
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.;
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat
ok, C.; Schlueter, T.; Smoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A>Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AE1139
A>Status: preliminary
A:Molecule type: DNA
```


F:1204-1599/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F:1682-1953/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>
F:2680-2751/Domain: acyl carrier protein homology <ACP2>
F:2804-3198/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
F:3295-3569/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
F:4320-4391/Domain: acyl carrier protein homology <ACP3>
F:4435-4830/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
F:5903-5974/Domain: acyl carrier protein homology <ACP4>
F:6018-6412/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
F:6513-6785/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT3>
F:7450-7521/Domain: acyl carrier protein homology <ACP5>

Query Match 5.9%; Score 116.5; DB 2; Length 7576;
Best Local Similarity 24.8%; Pred. No. 6.6;
Matches 85; Conservative 27; Mismatches 106; Indels 125; Gaps 18;

Qy 8 VTVLLGGDVMVLRGVQDILPHPGKPOLRERYMRDATGVRLAERVNGVRIPVDWRWPG 67
Db 5316 VTVAAPEGD--GRRVAVHARPGTGEWH-----ATGILARVARVASAVP---ETPSPWP 5366
Qy 68 EALAVLENTATDVCLINLETTITADGSEFADRPKVCYRMHPDNVPALTRALRHVC----- 121
Db 5367 PATA-----RPEDV-----GELADR--LARAGHTDG--PALP--RPRAAWADDDA 5405
Qy 122 -----ALANNHILDFGVQGLTDTVAALAGAGIQSVGAGADLLAA----- 160
Db 5406 VHAEVALADEQHAADERYGLH---PALLGAALALAGEGADLPSPADDDVRVHATGATTVRV 5462
Qy 161 -----RRSALTVTGHERR---VTGVSV-----AAESSGVPESWAARRDRPG 198
Db 5463 AVTATGIHLADETGPSVATGVARRRPLTVEGAVPGLLRPLDAIEAELPPPTATTGG--- 5519
Qy 199 VWLIRDPQAQDV-----ADVAAQVLA-----DKRPGDIAIVSMHWGS 236
Db 5520 ---LDDPVVPDVVILPAHGTGGGLGETRDLGARVALARSLFLLTDDRVD--AVLAVHTGP 5575
Qy 237 NWGYATAPGDVAFAPHLRIDAGIDMVHGSHHPRPIEYRKCP 279
Db 5576 GLAPAAAAG-----LVRTAQAEHPRIIVVDAEP 5604

RESULT 11
B8117
hypothetical protein NMB1147, NMB1185 [imported] - Neisseria meningitidis (strain MC58
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: B81117; G8112
R:Tetteelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: B81117
A:Molecule type: DNA
A:Residues: 1-211 <RET>
A:Cross-references: UNIPROT:Q9J513; GB:AE002463; GB:AE002098; NID:g7226377; PIDN:AAF4153
A:Experimental source: serogroup B, strain MC58
A:Accession: G81112
A:Molecule type: DNA
A:Residues: 1-211 <TED>
A:Cross-references: GB:AE002466; GB:AE002098; NID:g7226415; PIDN:AAF41569.1; PID:g722642
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1147; NMB1105

Query Match 5.8%; Score 116; DB 2; Length 211;
Best Local Similarity 25.1%; Pred. No. 0.092;
Matches 48; Conservative 29; Mismatches 66; Indels 48; Gaps 9;

Qy 214 VAAQVLADKRPGDIAIVSMHWGSNWGYAT-APGD-----VAFAPHLRIDAGIDM 260

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:24:25 ; Search time 85.1176 Seconds
(without alignments)
2286.133 Million cell updates/sec

Title: US-10-617-038-6
Perfect score: 1991
Sequence: 1 VAGNPDVTVTLGGDVMILGR.....VTRPDNLLEVPAANLTSKE 380

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1988	99.8	380	2	053771	O53771 mycobacteri
2	1982	99.5	380	2	Q7U1R0	Q7U1R0 mycobacteri
3	893	44.9	381	2	Q6AKV0	Q6AKV0 desulfotale
4	891	44.8	377	2	Q63LB1	Q63LB1 burkholderi
5	817.5	41.1	361	2	Q8T376	Q8T376 meloidogyne
6	771	38.7	378	2	Q8TX0	Q8TX0 methanosarc
7	337.5	17.0	405	2	Q93NC6	Q93NC6 myxococcus
8	300	15.1	351	2	Q74H56	Q74H56 geobacter s
9	277	13.9	516	2	Q67S32	Q67S32 symbiobacte
10	258	13.0	467	2	Q8XGR6	Q8XGR6 bacillus l
11	252	12.7	380	2	Q9AJM1	Q9AJM1 bacillus su
12	252	12.7	380	2	Q96738	Q96738 bacillus su
13	250.5	12.6	409	2	Q72QJ9	Q72QJ9 leptospira
14	249.5	12.5	409	2	Q8F5K7	Q8F5K7 leptospira
15	245.5	12.3	381	2	Q82SN6	Q82SN6 nitrosomona
16	240.5	12.1	388	2	Q67KZ9	Q67KZ9 symbiobacte
17	234	11.8	374	2	Q9A7U4	Q9A7U4 caulobacter
18	232.5	11.7	411	1	CAPA_BACAN	P19579 bacillus an
19	228.5	11.5	384	2	Q9RCW9	Q9RCW9 streptomyce
20	226.5	11.4	389	2	Q6SE61	Q6SE61 bacillus li
21	212.5	10.7	507	2	Q8DHZ2	Q8DHZ2 synecococc
22	204	10.2	461	1	Y4UA_RHISN	Q53195 rhizobium s
23	202.5	10.2	342	2	Q72PU0	Q72PU0 leptospira
24	202.5	10.2	342	2	Q8F6F6	Q8F6F6 leptospira
25	202	10.1	461	1	Y4SH_RHISN	P55651 rhizobium s
26	200	10.0	382	2	Q826A0	Q826A0 streptomyce
27	198.5	10.0	411	2	Q92VT6	Q92VT6 rhizobium m
28	197	9.9	102	2	Q6AKV3	Q6AKV3 desulfotale
29	194	9.7	297	2	Q7N4W8	Q7N4W8 photorhabdu
30	193.5	9.7	367	2	Q6HK98	Q6HK98 bacillus th
31	191.5	9.6	367	2	Q63CV1	Q63CV1 bacillus ce

32	190.5	9.6	367	2	Q81S24	Q81S24 bacillus an
33	188.5	9.5	367	2	Q9XB11	Q9XB11 bacillus ce
34	188.5	9.5	367	2	Q73A45	Q73A45 bacillus ce
35	188	9.4	367	2	Q81F23	Q81F23 bacillus ce
36	183	9.2	372	2	Q8YAU6	Q8YAU6 listeria mo
37	182.5	9.2	413	2	Q8A811	Q8A811 propionibac
38	182	9.1	372	2	Q72SF2	Q72SF2 listeria mo
39	177	8.9	476	2	Q8Y9L1	Q8Y9L1 listeria mo
40	176	8.8	424	2	Q9KL77	Q9KL77 rhizobium e
41	175	8.8	476	2	Q723D4	Q723D4 listeria mo
42	172	8.6	367	2	Q8R6T2	Q8R6T2 thermoaer
43	170.5	8.6	373	2	Q98I08	Q98I08 rhizobium l
44	166	8.3	475	2	Q92EE4	Q92EE4 listeria in
45	165	8.3	447	2	Q8G428	Q8G428 bifidobacte

ALIGNMENTS

RESULT 1
O53771 PRELIMINARY; PRT; 380 AA.
ID O53771 Q7D9L9;
AC O53771 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-1998 (TREMBLrel. 28, Last annotation update)
DE Hypothetical protein (CapA-related protein).
GN OrderedLocusNames=MT0602, RV0574C;
OS Mycobacterium tuberculosis
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=96344230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
Harris D.E., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III,
Tekai F., Badcock K., Basham D., Brown D., Chillingworth T.,
Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J.D., Deboy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
DR EMBL; BX842573; CAA17445.1; -;
DR EMBL; AE000516; AAK44825.1; -;
DR PIR; E70933; E70933.
DR TIGR; MT0602; -;
DR TubercuList; RV0574C; -;
SQ SEQUENCE 380 AA; 41547 NW; 5197A2C71A981EF6 CRC64;

Query Match 99.8%; Score 1988; DB 2; Length 380;
Best Local Similarity 99.7%; Pred. No. 3e-142;
Matches 379; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAGNPDVTVTLGGDVMILGRGVQDQILPHPGKPLQRERYMRDATGYVRLAERVNGRIPLPV 60


```
Db 371 NR 373
RESULT 4
Q63LB1 PRELIMINARY; PRT; 377 AA.
ID Q63LB1
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Putative capsule biosynthesis protein.
GN ORENames=BPSS1097;
OS Burkholderia pseudomallei K96243.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=272560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K96243;
RX PubMed=15377794;
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdano-Tarraga A.M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crosset B., Davis P., DeShazer D.,
RA Feltwell T., Fraser A., Hance Z., Hauser H., Holooyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinovitch E., Rutherford K., Sanders M., Simmonds M.,
RA Songvilai S., Stevens K., Tumapa S., Vesaratchavest M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RT Burkholderia pseudomallei.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR EMBL; BX571966; CAH38565.1; -.
SQ SEQUENCE 377 AA; 41534 MW; 19640BE373B31C07 CRC64;

Query Match 44.8%; Score 891; DB 2; Length 377;
Best Local Similarity 51.0%; Pred. No. 3.3e-59;
Matches 181; Conservative 50; Mismatches 120; Indels 4; Gaps 4;

QY 8 VTVLLGDMVGRGVQILPHGPKPOLRERYMDATGYVRLAERVNGRIPLPVDWNPWG 67
Db 13 VRFLCGDVTGRIQILPHGPKPOLRERYMDATGYVRLAERVNGRIPLPVDWNPWG 72
QY 68 EALAVLENTATDVCLINLETTI-TADGEFADRPKPCVYRMHPDNPALTALRPHVCALANN 126
Db 73 DALAELDRVPHRVIYNLETTITADGEFADRPKPCVYRMHPDNPALTALRPHVCALANN 131
QY 127 HILDFGVQGLTDTVAALAGAGIQSVGAGADLLAARRSALVTGHERRVIVGSVAESSGV 186
Db 132 HSLDWGREGLADTLRLRAGIRTAGAGDDAARPATLGNVAGRRVLVYAYAAETSGV 191
QY 187 PESWAARRDRPGVWLIRDPADRDVADVAQVLADKRPDIATVSMHWSNWCYATAPGD 246
Db 192 PPSWAATPRGGVNYLAD-ISSGRATQIGERIAQRREGDLVSVLHWSNWCYATAPGD 250
QY 247 VAFARHLIDAG-IDMVHGHSSHPRIETIYRGKPILYGCGDVVDYEGIGGHSFSELR 305
Db 251 HAFARHLIDTGAADIVVGHSSHVKGIEIYRERLILYGGCDLNDYEGIGHQPPFDLA 310
QY 306 LLYLTVTDPASGNLISQMLPLRVSRMLQRASQDTTEWLNTIERSRFRGVTRP 360
Db 311 LMYFPVLDAGSALVELSAVPMQIRNLRLQAPADGKAWLHAYLVERSRFRGVTH 365

RESULT 5
ID Q8T376 PRELIMINARY; PRT; 361 AA.
AC Q8T376
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Polyglutamate synthase.

GN Name=ths;
OS Meloidogyne artiellia.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
OC Tylenchoidae; Heteroderidae; Meloidogyninae; Meloidogyne.
OX NCBI_TaxID=42426;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21585544; PubMed=11728474; DOI=10.1016/S0014-5793(01)03132-5;
RA Veronico P., Jones J.T., Di Vito M., De Giorgi C.;
RT "Horizontal transfer of a bacterial gene involved in polyglutamate
RT biosynthesis to the plant-parasitic nematode Meloidogyne artiellia.";
RL FEBS Lett. 508:470-474(2001).
DR EMBL; AJ310212; CAC84452.1; -.
SQ SEQUENCE 361 AA; 40486 MW; 8939EFABCD3B48C CRC64;

Query Match 41.1%; Score 817.5; DB 2; Length 361;
Best Local Similarity 44.5%; Pred. No. 1.1e-53;
Matches 159; Conservative 69; Mismatches 124; Indels 5; Gaps 4;

QY 17 MLGRGVQILPHGPKPOLRERYMDATGYVRLAERVNGRIP-LPVDWNPWGALAVLEN 75
Db 1 MLGRGIDQILPHGPNRPFESCVKNANIYVRLAEQTSGLPKRPLHYAVPWGTLAELEE 60
QY 76 TATDVCLINLETTITADGEFADRPKPCVYRMHPDNPALTALRPHVCALANNHILDFGYQG 135
Db 61 HRPDARIINLETSVTTSSESSWPAKGIHYRMHPNVECTISARVDDCCVLNNHVDLGHAG 120
QY 136 LTDVVAALAGAGIQSVGAGADLLAARRSALVTGHERRVIVGSVAESSGVPSWAARRD 195
Db 121 LPETLETLHSSGKTAGAGINENDASAPALMESEGRLLIFGYADESSGVPSWAATKR 180
QY 196 RFGVWLIRDPADRDVADVAQVLADKRP--GDIAIVSMHWSNWCYATAPGDVAFARHL 253
Db 181 REGVNFLLHDGISAASADSICRQIAKERPERDEKDIVICSLHWSNWCYQISREERQFAHRL 240
QY 254 IDAG-IDMVHGHSSHPRIETIYRGKPILYGCGDVVDYEGIGGHSFSELRLLYLTVT 312
Db 241 IDSGCVDIHGHSSHHAKGLEIYRKLILYGGDFLNDYEGISGHDEYRDDCTLMTFPRV 300
QY 313 DPASGNLISQMLPLRVSRMLQRASQDTTEWLNTIERSRFRGVTRP-DNLL 368
Db 301 NSSDGTLLSLFAPVPMQIKHMLNRANDEVGILSNVMQRECAKFEVNVQRDPVKNIL 357

RESULT 6
Q8TJX0 PRELIMINARY; PRT; 378 AA.
ID Q8TJX0
AC Q8TJX0
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein MA3657.
GN OrderedLocusNames=MA3657;
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238; DOI=10.1101/gr.223902;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA Fitzhugh W.E., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Taylor J., Stange-Thomann N., DeArellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Kazycki J.A.,
RA Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Perry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.T.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
```

```

RT and physiological diversity."
RL EMBL: AE011074; RA007012.1; -.
DR Complete proteome.
KW SEQUENCE 378 AA; 42652 MW; DF35B99FF23012B CRC64;
Query Match 38.7%; Score 771; DB 2; Length 378;
Best Local Similarity 41.3%; Pred. No. 4e-50;
Matches 149; Conservative 73; Mismatches 137; Indels 2; Gaps 2;
QY 7 VTVLLGGDVMLGRGVQDQILPHPGKPOLRERYMRDATGYVRLAERVNGRIPLPVDWRMPW 66
Db IISIFLCGDVMPGRGIDQILPHPGDPEIVEPCMSAKWVYTAEKNGEINKPVDVSYVM 67
QY 67 GEALAVENTATDVCLINLETTITADGEFADRPKPCYCRMHPDNVPALTALRHVICALANN 126
Db GDALKEFNKKPDLKLIINLETSITSSRDYWKHEVHYKLNPNQNDIVIKNVGIDFTSLANN 127
QY 127 HILDFGYOGLTDTVAALAGAGIQSVGAGADLLAARRSALVTVGHERRVIVGSVAESSGV 186
Db HVLWDGYSKGLDETAKIKFAGAGIDKSLAEKIAVFDLKGKRVLIYSCGSITSGI 187
QY 187 PESWAARRDRPGVWLIRDPADRDVAAQVLADKRPDIAIVSMHMGNSNGYATAPGD 246
Db PNEWSAENKAGVNLIPD-YSIDTVRNIEKPDIAIRGENDISIVSIHWGNGWYKIPGAH 246
QY 247 VAFARLID-AGIDMVHGHSHHPRIEYRGKPIYKGGDVGDDVDDYEGIGGHESFRSEL 305
Db MEFAHNLVDLAGVNTVHGHSHHPRIEYVKNRLIYKGGDFINDYEGITDCMYREDLA 306
QY 306 LLYLTVDTPASGNLISLQMLPLRVSRMRLOASQDTEWLRTIERISRRRGIRVVTRPD 365
Db LMYFAGFDADTGNISKLVPMQIKFNLVAGHNDABWTKNILNREGKADTFEKLUSD 366
QY 366 N 366
Db 367 N 367

RESULT 7
QJ3NC6 PRELIMINARY; PRT; 405 AA.
ID QJ3NC6
AC QJ3NC6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Polyglutamate synthase CapA.
GN Name=capA;
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cystobacterineae; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
RN SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC STRAIN=DZF1;
RA Nariya H., Inouye S.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL: AF377339; AAK64444.1; -.
SQ SEQUENCE 405 AA; 43528 MW; 3E96B3474E0004E7 CRC64;

Query Match 17.0%; Score 337.5; DB 2; Length 405;
Best Local Similarity 32.6%; Pred. No. 2.8e-17;
Matches 107; Conservative 44; Mismatches 138; Indels 39; Gaps 10;
QY 8 VTVLLGGDVMLGRGVQDQILPHPGKPOLRERYMRDATGYVRLAERVNGRIPLPVDWRMPW 67
Db VSIIVGGDVTLGHNYQ-----THYD--AETAAGR---PQDALLAYG 95
QY 68 EALAVENTATDVCLINLETTITADGEFADRPKPCYCRMHPDNVPALTALRHVICALANNH 127
Db FOEVKPLTDAADLFVFNUECCEFTERGEKLP-KNFNFRAPKPELVGLTLAGGVDVWSLANNH 154
QY 128 ILDFGYOGLTDTVAALAGAGIQSVGAGADLLAARRSALVTVGHERRVIVGSVAESSGV- 186

and physiological diversity."
RL EMBL: AE011074; RA007012.1; -.
DR Complete proteome.
KW SEQUENCE 378 AA; 42652 MW; DF35B99FF23012B CRC64;
Query Match 38.7%; Score 771; DB 2; Length 378;
Best Local Similarity 41.3%; Pred. No. 4e-50;
Matches 149; Conservative 73; Mismatches 137; Indels 2; Gaps 2;
QY 7 VTVLLGGDVMLGRGVQDQILPHPGKPOLRERYMRDATGYVRLAERVNGRIPLPVDWRMPW 66
Db IISIFLCGDVMPGRGIDQILPHPGDPEIVEPCMSAKWVYTAEKNGEINKPVDVSYVM 67
QY 67 GEALAVENTATDVCLINLETTITADGEFADRPKPCYCRMHPDNVPALTALRHVICALANN 126
Db GDALKEFNKKPDLKLIINLETSITSSRDYWKHEVHYKLNPNQNDIVIKNVGIDFTSLANN 127
QY 127 HILDFGYOGLTDTVAALAGAGIQSVGAGADLLAARRSALVTVGHERRVIVGSVAESSGV 186
Db HVLWDGYSKGLDETAKIKFAGAGIDKSLAEKIAVFDLKGKRVLIYSCGSITSGI 187
QY 187 PESWAARRDRPGVWLIRDPADRDVAAQVLADKRPDIAIVSMHMGNSNGYATAPGD 246
Db PNEWSAENKAGVNLIPD-YSIDTVRNIEKPDIAIRGENDISIVSIHWGNGWYKIPGAH 246
QY 247 VAFARLID-AGIDMVHGHSHHPRIEYRGKPIYKGGDVGDDVDDYEGIGGHESFRSEL 305
Db MEFAHNLVDLAGVNTVHGHSHHPRIEYVKNRLIYKGGDFINDYEGITDCMYREDLA 306
QY 306 LLYLTVDTPASGNLISLQMLPLRVSRMRLOASQDTEWLRTIERISRRRGIRVVTRPD 365
Db LMYFAGFDADTGNISKLVPMQIKFNLVAGHNDABWTKNILNREGKADTFEKLUSD 366
QY 366 N 366
Db 367 N 367

RESULT 8
Q74H56 PRELIMINARY; PRT; 351 AA.
ID Q74H56
AC Q74H56;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Capsule biosynthesis protein, putative.
GN OrderedLocusNames=GSU0036;
OS Geobacter sulfurreducens.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
OC Geobacteraceae; Geobacter.
OX NCBI_TaxID=35554;
RN SEQUENCE FROM N.A.
RC STRAIN=PCA / ATCC 51573;
RX PubMed=14671304; DOI=10.1126/science.1088727;
RA Methe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,
RA Heidelberg J.F., Wu D., Wu M., Ward N.L., Beanan M.J., Dodson R.J.,
RA Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,
RA Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,
RA Daviden T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,
RA Weidman J.F., Khouli H.M., Feldblyum T.V., Utterback T.R.,
RA Van Aken S.E., Lovley D.R., Fraser C.M.;
RT "Genome of Geobacter sulfurreducens: metal reduction in subsurface
RT environments."
RL Science 302:1967-1969 (2003).
DR EMBL: AE017180; AAR33371.1; -.
DR TIGR; GSU0036; -.
KW Complete proteome.
SQ SEQUENCE 351 AA; 37570 MW; 24E7FEEA3B39E077 CRC64;

Query Match 15.1%; Score 300; DB 2; Length 351;
Best Local Similarity 30.7%; Pred. No. 1.7e-14;
Matches 107; Conservative 37; Mismatches 138; Indels 66; Gaps 12;
QY 14 GDVNL-GRGVQDQILPHPGKPOLRERYMRDATGYVRLAERVNGRIPLPVDWRMPGEALAV 72
Db GDVNLGRGVQDQILPHPGKPOLRERYMRDATGYVRLAERVNGRIPLPVDWRMPGEALAV 72
QY 73 LENTATDVCLINLETTITADGEFADRPKPCYCRMHPDNVPALTALRHVICALANNHILDPG 132
Db LRR--SGLVGNLEAPLTDGGREFRAKRFRRFPAVAAAALAKRAGFSYMTLANNHMDFG 117
QY 133 YQGLTDTVAALAGAGIQSVGAGADLLAARRSALVTVGHERRVIVGSVAESSGVPESWAA 192
Db YQGLTDTVAALAGAGIQSVGAGADLLAARRSALVTVGHERRVIVGSVAESSGVPESWAA 192
QY 118 ADGLSDTIHLNRRNGIAFAGAGPSIADARRASVTV----RGQTVAFVLAFLYSLTQIEPFA 173
Db ADGLSDTIHLNRRNGIAFAGAGPSIADARRASVTV----RGQTVAFVLAFLYSLTQIEPFA 173
QY 193 RRDRPGVNLIRDPADRDVAAQVLADKRPDIAIVSMHMGNSNGYATAPGDVA 248
Db TEGRPG-----TAPGYAGHYLADIRRVSSADHVVVSHWGQERAAALPSYQIE 222
QY 249 FAHRLIDAGIDMVHGHSHHPRIEYRGKPIYKGGDVGDDVDDYEGIGGHESFRSELRLLY 308
Db TAHRADAGADIVIGHHHPVQLQGIIEYRGSPFIYSLGNFA---FGSRSPSADRSTIARVT 279
QY 309 LTVTDPASGNLISLQMLPLRV--SRMRLOASQDTEWLRTIERISR 354

```


[illegible]


```

RC STRAIN=Fiocruz L1-130;
RX PubMed=15028702; DOI=10.1128/JB.186.7.2164-2172.2004;
RA Nascimento A.L.T.O., Ko A.I., Martins E.A.L., Monteiro-Vitorello C.B.,
RA Ho P.L., Haake K.T.O., Verjovski-Almeida S., Hartkeerl R.A.,
RA Marques M.V., Oliveira M.C., Mencia C.F.M., Leite L.C.C., Carter H.,
RA Coutinho L.L., Degraive W.M., Dellagostin O.A., El-Dorry H.,
RA Ferro E.S., Ferro M.I.T., Furlan L.R., Gamberini M., Gigliotti E.A.,
RA Goes-Neto A., Goldman G.H., Goldman M.H.S., Harakava R.,
RA Jeronimo S.M.B., Junqueira-de-Azevedo I.L.M., Kimura E.T.,
RA Kuranee E.E., Lemos E.G.M., Lemos M.V.F., Marino C.L., Nunes L.R.,
RA de Oliveira R.C., Pereira G.G., Reis M.S., Schrieffer A.,
RA Siqueira W.J., Sommer P., Teai S.M., Simpson A.J.G., Ferro J.A.,
RA Camargo L.E.A., Kitajima J.P., Setubal J.C., Van Sluys M.A.;
RA "Comparative genomics of two Leptospira interrogans serovars reveals
RT novel insights into physiology and pathogenesis.";
RL J. Bacteriol. 186:2164-2172(2004).
DR EMBL; AB017295; AAS70685.1; -.
KW Complete proteome.
SQ SEQUENCE 409 AA; 45899 MW; 812AC93FB40C668D CRC64;

Query Match 12.6%; Score 250.5; DB 2; Length 409;
Best Local Similarity 27.1%; Pred. No. 1.1e-10;
Matches 88; Conservative 49; Mismatches 139; Indels 49; Gaps 10;

QY 4 NPDVTVLLGGVMDLGRGVQDQLPHPGKQPLRERYMDATGYVRLAERNVGRIPLPVDWR 63
DB 71 DPGLLLKLLAGGDMFNGI-----RDTI-----RXHGEI-APV--- 102
QY 64 WPMGEALAVLENTATDVCLINLETTITADGEFADRPKPCVYRHPDNPVLTALRPHVCAL 123
DB 103 -----EGLRPLFNEA-DFRMINLETPIVASKTEESKKAYIFTAHEKDLNLSKYLGVDMVFL 157
QY 124 ANNHILDGFGYQGLTDTVAALAGAGIOSVGAGADLLAARRSALVTYV-GHERRVIVGSVAE 182
DB 158 GNNHSPDHGQOGMNETDILNKNILFVGAGKKLPEVLEPLNLNKGSLDLRIHVSVAIAE 217
QY 183 SSGVPESWAARRDRPGVWLIRDPADRDVADVAOVLADKRPDGIATVSMHWSNGYAT 242
DB 218 -----QTHVATATRSVAPFLPESLR--ATFEKNILPKDHTPPVRIVSLHWGVEYSPFP 270
QY 243 APGDVAFARHLIDAGIDMVHGHSSHPRIETIYRGKPILYGCGDVVDVDEGGHESFPS 302
DB 271 TSDQRIARSLDSGVKIVIGHHPHIPQGVVEYKGGVILYSLGNLI-----FGSRNSYLN 325
QY 303 ELRLYLTVTPASGNLISLQMLPL 327
DB 326 HNLIAILHI---RNNHLRVELVPI 347

RESULT 15
Q82SN6 PRELIMINARY; PRT; 381 AA.
ID Q82SN6;
AC Q82SN6;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Possible poly-gamma-glutamate synthesis protein.
GN Name=pgsA,ywtB; OrderedLocusNames=NE2271;
OS Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.
OX NCBI_TaxID=915;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19718 / IFO 14298;
RX MEDLINE=22586410; PubMed=12700255;
RX DOI=10.1128/JB.185.9.2759-2773.2003;
RA Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.L.,
RA Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
RA Arciero D.M., Hommes N.G., Whittaker M.M., Arp D.J.;
RT "Complete genome sequence of the ammonia-oxidizing bacterium and
RT obligate chemolithoautotroph Nitrosomonas europaea.";
RL J. Bacteriol. 185:2759-2773(2003).
DR EMBL; BX321864; CAD86183.1; -.
KW Complete proteome.
SQ SEQUENCE 381 AA; 42389 MW; 06598A77CAB47125 CRC64;

Query Match 12.3%; Score 245.5; DB 2; Length 381;
Best Local Similarity 28.2%; Pred. No. 2.4e-10;
Matches 87; Conservative 47; Mismatches 119; Indels 55; Gaps 13;

QY 67 GEAL---AVLENTATDVCLINLETTITADGEFADRPKPCVYRHPDNPVLTALRPH---- 119
DB 24 GSAILGNAIEIKTADFVNECEPLT-DHQAINK-----SGPALRA-GPOCAGS 72
QY 120 -----VCALANNHILDGFGYQGLTDTVAALAGAGIOSVGAGADLLAARRSALVTYVGHERRV 174
RT "Unique physiological and pathogenic features of Leptospira

```

```
Db 73 IADFTVAGLANNHSLDYGQGLIDTITACRSVGVSTVGAGINLAEQAQIHIISKVKGKLA 132
Qy 175 IVGSVAABESSGVPESWAARRDPGVWLIRDPAQRDVADDVAAQVLAADKRPGLIAIVSMHW 234
Db 133 VIAVAEHFN-----QSENNGFG-----SAPLDPVDNY-YQIREAQAKADIVIVTIHG 179
Qy 235 GSNMGYATAPGDVAFAPHRLLIDAGIDMVHGHSHHPRPIEIVRGKPILYCGGDVDDYEGI 294
Db 180 GNEHFHYRPGRLKLCCKHYIDLGVNAVICHHPHVEGAYEINYGRPIVYSLGNFVFDTLSM 239
Qy 295 GGHE---SFRSELRLLYLTVTDPASGNLISLQMLELR-----VSRMRLORASQTD-----T 342
Db 240 -VHEWDVGYMAKLKFEVDCTFEA-----IEIIPYQSIITVEGVELLRGDERDKAVSKI 292
Qy 343 EWLKNTIE 350
Db 293 EALRNAVQ 300
```

Search completed: October 15, 2005, 04:35:54
Job time : 87.1176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:20:35 ; Search time 105.54 Seconds
(without alignments)
1454.845 Million cell updates/sec

Title: US-10-617-038-7
Perfect score: 1999
Sequence: 1 VTHVREADANIDLLGLD.....ATRTSGGAQQDDLAGSSP 397

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1999	100.0	397	8	Adi37286 M. tuberc
2	190.5	9.5	324	5	Abu05518 M. tuberc
3	168	8.4	185	7	Adj83099 Adenylate
4	162	8.1	316	5	Abu05827 M. tuberc
5	159	8.0	356	8	Ads22942 Bacterial
6	157	7.9	670	2	Aay14917 Extended
7	157	7.9	670	5	Abb73523 M vaccae
8	157	7.9	670	6	Abp70898 Mycobacte
9	157	7.9	722	2	Aay14915 Amino aci
10	157	7.9	722	5	Abb73521 M vaccae
11	151.5	7.6	566	4	Abb61743 Drosophil
12	151	7.6	419	2	Aay60149
13	151	7.6	419	2	Aay14896 M. vaccae
14	151	7.6	419	5	Abb73502 M vaccae
15	148.5	7.4	194	7	Adj83100 Adenyl
16	148.5	7.4	1690	4	Abb64305 Drosophil
17	148.5	7.4	1690	8	Adq99708 Antagonis
18	144.5	7.2	619	3	Aay51608 Human sol
19	144.5	7.2	619	7	Adc38455 Human pro
20	144.5	7.2	619	8	Adq18090 Human sof
21	136.5	6.8	9234	7	Adc26980 Sorangium
22	136	6.8	1029	2	Aar38862 GC-A 2/1
23	136	6.8	1037	7	Ades6805 Rat Prote
24	135	6.8	462	4	Aag81158 Mycobacte
25	133	6.7	1061	5	Aau11281 Human atr

Aau11280 Human atr
Adq99892 Antagonis
Abm80153 Tumour-as
Abb11783 Human ANP
Abb60360 Drosophil
Adr01215 Farneeyl
Abu32161 Protein e
Adn21040 Bacterial
Ab076858 Pseudomon
Abm83201 Human dia
Abm83200 Human dia
Abb9684 Drosophil
Adq99690 Antagonis
Abm83202 Human dia
Abd00055 Mycobacte
Aar10867 NPRB(Pro6
Aab94737 Human pro
Adj33095 Human ade
Adj82968 Human ade
Adq99906 Antagonis

ALIGNMENTS

RESULT 1

ADI37286
ID ADI37286 standard; protein; 397 AA.

XX ADI37286;

XX DT 22-APR-2004 (first entry)

XX DE M. tuberculosis low oxygen induced antigen Rv1264 SEQ ID NO:7.

XX KW mycobacterial infection; vaccine; tuberculosis;

XX KW Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;

XX KW low oxygen induced antigen.

OS Mycobacterium tuberculosis.

XX WO2004006952-A2.

XX PD 22-JAN-2004.

XX PF 08-JUL-2003; 2003WO-DK000477.

XX PR 13-JUL-2002; 2002DK-00001098.

XX (STAT-) STATENS SERUM INST.

XX Andersen P, Rosenktrands I, Stryhn A;

XX WPI; 2004-122778/12.

XX N-PSDB; ADI37331.

Use of one or more polypeptides or their fragments, which are expressed during the latent stage of the mycobacterial infection, and/or nucleic acids encoding the polypeptides, for a therapeutic vaccine against tuberculosis.

Claim 3; SEQ ID NO 7; 76pp; English.

The present invention describes polypeptides or their fragments, which are expressed during the latent stage of a mycobacterial infection, and/or nucleic acids encoding the polypeptides, which are useful for creating a therapeutic vaccine against tuberculosis. Also described: (1) a therapeutic vaccine against tuberculosis comprising one or more polypeptides; (2) a method for treating an animal, including a human being, with tuberculosis caused by virulent mycobacteria, e.g. by Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for immunising an animal, including a human being, against tuberculosis caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis

Db 420 ISGGDYRLALPVLRSDFGDLTTAFNDM-----SRNLSTKIDELLGEEERAENQRLMLSLMPE 476
 Qy 211 -----PGARQVTVAFADLVGFTQLGSEVSAEEL-----GHLAGRLAGLARDL 252
 Db 476 PVMQRYLDGSEETIAQDHKNVTVI FADMMGLDEL SRMLTSEELMVMVNDL TRQF DAAAE SL 535
 Qy 253 TAPPVWFIKITGDVAMLVCPDPAPLLDTVLKLV-----EVDV-----TDNNFPRLRAGV 301
 Db 536 ---GVDHVRTLHDGYLASCGLGVGRPLDNVRRTVNFAIEMDRIIDRHA AESG HDLRLRAGI 592
 Qy 302 ASGNMAYS-----RAGDWFSGPVNVASRV-TGVARPGAVLVADSVRE-----ALG 344
 Db 593 DTGSAASGLVGRSTFLAYDMWGSADVAVANQVGRGSPQFG-IYVTSRVHEVMQETLDFVAAG 651
 Qy 345 DAPEADGFQ--WSPAGPRR 361
 Db 652 EVUGERGVEVWRQLQGHRR 670

RESULT 7
 ABB73523
 ID ABB73523 standard; protein; 670 AA.
 XX AC ABB73523;
 XX DT DT
 XX 08-APR-2002 (first entry)
 DE M vaccae GV-38B truncated protein SEQ ID NO: 178.
 XX Skin disorder; psoriasis; atopic dermatitis; allergic contact dermatitis;
 KW alopecia areata; skin cancer; Mycobacterium vaccae; melanoma; cytostatic;
 KW antipsoriatic; dermatological; antiinflammatory; antiallergic;
 KW Th2 immune response; immunomodulatory.
 XX Mycobacterium vaccae.
 OS
 XX US6328978-B1.
 XX 11-DEC-2001.
 XX 02-JUN-1999; 99US-00324542.
 XX 23-DEC-1997; 97US-00997080.
 XX (GENE-) GENESIS RES & DEV CORP LTD.
 XX Watson JD, Tan PLJ, Prestidge R;
 PI
 XX WPI; 2002-138361/18.
 DR N-PSDB; ABL36287.
 XX
 XX Inhibiting skin inflammation associated with skin disorder e.g.
 PT psoriasis, by administering composition comprising delipidated and
 PT deglycolipidated Mycobacterium vaccae cells or Mycobacterium vaccae
 PT culture filtrate.
 XX
 XX Example 9; Col 195-200; 116pp; English.
 PS
 XX The present invention relates to a method of inhibiting skin inflammation
 CC associated with a skin disorder selected from psoriasis, atopic
 CC dermatitis and allergic contact dermatitis, which involves administering
 CC a composition containing delipidated and deglycolipidated Mycobacterium
 CC vaccae cells or M. vaccae culture filtrate. The skin disorder to be
 CC treated may also include alopecia areata, and skin cancers such as basal
 CC cell carcinoma, squamous cell carcinoma and melanoma. The composition
 CC acts by inhibiting the Th2 immune response. The present sequence is a
 CC protein described in the exemplification of the invention
 XX
 SQ Sequence 670 AA;

Query Match 7.9%; Score 157; DB 5; Length 670;
 Best Local Similarity 23.9%; Pred. No. 7.9e-06;
 Matches 105; Conservative 55; Mismatches 159; Indels 120; Gaps 20

XX 23-DEC-1997; 97US-00996624.
PR 23-DEC-1997; 97US-00997080.
PR 23-DEC-1997; 97US-00997362.
PR 11-JUN-1998; 98US-00095855.
PR 17-SEP-1998; 98US-00156181.
PR 04-DEC-1998; 98US-00205426.
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX Tan P, Watson J, Visser ES, Skinner MA, Prestidge RL;
PI WPI; 1999-430163/36.
DR N-PSDB; AAZ11379.
XX Enhancing immune response to an antigen.
PT
PS Claim 1; Page 222-224; 243pp; English.
XX The invention provides heat-killed Mycobacterium vaccae, or recombinant
CC M. vaccae proteins. The M. vaccae proteins may be employed to activate T
CC cells and natural killer cells, to stimulate the production of cytokines,
CC to enhance the expression of co-stimulatory molecules on dendritic cells
CC and monocytes, and to enhance dendritic cell maturation and function. The
CC proteins can be expressed by standard recombinant methodology.
CC Pharmaceutical compositions comprising the proteins or nucleic acid
CC sequences encoding the proteins can be used for the treatment,
CC prevention, and detection of disorders including infectious diseases,
CC immune disorders and cancer. In particular, the compounds and methods are
CC used for treatment of diseases of the respiratory system, such as
CC mycobacterial infections, asthma, allergies, tuberculosis, leprosy,
CC sarcoidosis and lung cancers, and disorders of the skin such as
CC psoriasis, atopic dermatitis, eczema, allergic contact dermatitis,
CC alopecia areata, and skin cancers such as basal carcinoma, squamous cell
CC carcinoma and melanoma
XX
SQ Sequence 722 AA;
Query Match 7.9%; Score 157; DB 2; Length 722;
Best Local Similarity 23.9%; Pred. No. 8.8e-06;
Matches 105; Conservative 55; Mismatches 159; Indels 120; Gaps 20;
QY 14 DLLGLDGLGTA-----RAERAKLVEWLLFGGTPDEI-----RATNPP 51
Db 313 DTGMDGTGTILVGPDLNMRSDRLFRENRKFLADWEGG-TPPEVADESVDRRGTTLV 371
QY 52 LLLATRLHV-----GDDGT-----YVSAREISENYGVDELLELQVRAVGLARVDDPDVAV 102
Db 372 QPVTTRSVEAQRGNTGTTIEDDYLGHLEALQAVSPVDLPGLHWI-----VAKID----- 421
QY 103 HNRADGEAAARQRFVELGLNPDQVVLVVRVLAEGLSHAAEAMRYTALFAIMRPGATELD 162
Db 422 ----TDEAFAPVAQFTR-----TLVSTVLIIFGVSLAAMLLARLFFVPIRLQAGAOQ 471
QY 211 -----PGARQVTVAFADLVGFTQLGEVVSAEEL-----CHLAGRLAGLARDL 252
Db 528 FVNRVLDGEETIAQDHKNVTFIFADMGMGLDELRLMTSEELMVVNDLTQFDDAAAE 587
QY 253 TAPVNVFIKTIGDAVMLVCPDPAPLDTVLKLV-----EVVD-----TDNNFPRLRAGV 301
Db 588 ---GVDHVRTLHDGYLASCGLGVPLDNVRRVTNFAIEMDRIIDRHAESGHDRLRAGI 644
QY 302 ASGMAYS-----RAGDWFGSPVNVASRY-TGVARPGAVLVADSVRE-----ALG 344
Db 645 DTGSAASGLVGRSTLAYDMWGSADVAVANQVQSGPQFG-IYVTSRVHEVNETLDFVAAAG 703
QY 345 DAPEDAGFQ--WSFAGPFR 361
Db 704 EVVGERGVETVWRLOQHRR 722

RESULT 10

ABE73521
ID ABB73521 standard; protein; 722 AA.
XX
AC ABB73521;
XX
DT 08-APR-2002 (first entry)
XX
DE M vaccae GV-38B protein SEQ ID NO: 174.
XX
KW Skin disorder; psoriasis; atopic dermatitis; allergic contact dermatitis;
KW alopecia areata; skin cancer; Mycobacterium vaccae; melanoma; cytostatic;
KW antipsoriatic; dermatological; antiinflammatory; antiallergic;
KW Th2 immune response; immunomodulatory.
XX
OS Mycobacterium vaccae.
XX
PN US6328978-B1.
XX
PD 11-DEC-2001.
XX
PF 02-JUN-1999; 99US-00324542.
XX
PR 23-DEC-1997; 97US-00997080.
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI Watson JD, Tan PJ, Prestidge R;
XX
DR WPI; 2002-138361/18.
DR N-PSDB; ABL36285.
XX
PT Inhibiting skin inflammation associated with skin disorder e.g.
PT psoriasis, by administering composition comprising delipidated and
PT deglycolipidated Mycobacterium vaccae cells or Mycobacterium vaccae
PT culture filtrate.
XX
PS Example 9; Col 187-192; 116pp; English.
XX
CC The present invention relates to a method of inhibiting skin inflammation
CC associated with a skin disorder selected from psoriasis, atopic
CC dermatitis and allergic contact dermatitis, which involves administering
CC a composition containing delipidated and deglycolipidated Mycobacterium
CC vaccae cells or M. vaccae culture filtrate. The skin disorder to be
CC treated may also include alopecia areata, and skin cancers such as basal
CC cell carcinoma, squamous cell carcinoma and melanoma. The composition
CC acts by inhibiting the Th2 immune response. The present sequence is a
CC protein described in the exemplification of the invention
XX
SQ Sequence 722 AA;
Query Match 7.9%; Score 157; DB 5; Length 722;
Best Local Similarity 23.9%; Pred. No. 8.8e-06;
Matches 105; Conservative 55; Mismatches 159; Indels 120; Gaps 20;
QY 14 DLLGLDGLGTA-----RAERAKLVEWLLFGGTPDEI-----RATNPP 51
Db 313 DTGMDGTGTILVGPDLNMRSDRLFRENRKFLADWEGG-TPPEVADESVDRRGTTLV 371
QY 52 LLLATRLHV-----GDDGT-----YVSAREISENYGVDELLELQVRAVGLARVDDPDVAV 102
Db 372 QPVTTRSVEAQRGNTGTTIEDDYLGHLEALQAVSPVDLPGLHWI-----VAKID----- 421
QY 103 HNRADGEAAARQRFVELGLNPDQVVLVVRVLAEGLSHAAEAMRYTALFAIMRPGATELD 162
Db 422 ----TDEAFAPVAQFTR-----TLVSTVLIIFGVSLAAMLLARLFFVPIRLQAGAOQ 471
QY 163 IAKGSOALVSQIVPL-----LGPMDQLMFLMQLRHHMETEAVNAGERAAGKPL----- 210
Db 472 ISGGDYRLALPVLRSRDEFGLDTAFNDM-----SRNLSIKDELGGERAENQRLMLSLMPE 527

CC The products also have the ability to induce cell proliferation and
CC cytokine production (e.g. interferon-gamma and interleukin-12 production)
CC in T cells, NK cells, B cells, or macrophages. They can be used for
CC enhancing immune responses for use in vaccines or immunotherapy of
CC infectious diseases and cancers
XX
SQ Sequence 419 AA;

Query Match 7.6%; Score 151; DB 2; Length 419;
Best Local Similarity 23.3%; Pred. No. 1.6e-05;
Matches 103; Conservative 60; Mismatches 163; Indels 116; Gaps 19;
QY 14 DDLGLDGLGTA-----RAERAKLVLEWLLLEQGITPDEI-----RATNPP 51
DB 11 DTGMDGTETILVGPDLNMRSDSRFLFRENREKFLADVVEGG-TPEVADESVDRRGTTLV 69
QY 52 LLLATRLHV---GDDGT-----YVSAREISENYGVDLLELQVQRAVGLARVDDPAVV 102
DB 70 QPVTTSRVEEAQRNGTGTIEDDYLGEALQAYSPVDLPGLHWI-----VAKID----- 119
QY 103 HMRADGEAAARQRFVELGLNPDQVVLVVRVLAEGLSHAAEAAMRYTALRAIMRPGATELD 162
DB 120 ----TDEAFAPVAQFTR-----TLVLTSTVIIIFGVSLAAMLLARLFVVRPIRLQAGAQ 169
QY 163 IAKGSOALVSQIVPL-----LGPMDQLFMQLRHMMETEAVNAGERAAKGKPL----- 210
DB 170 ISGGDYRLALPVLRSRDEFGLTAFNDM-----SRNLSIKDELLEGERAENQRLMSLMPE 225
QY 211 ----PGARQVTVAFADLVGFTOLGEVVSABEL-----GHLAGLAGLARDL 252
DB 226 PVMQRYLDEETIAQDHKNVTIFADMMGLDLSRLTSEELMVVNDLTQFDAAAESL 285
QY 302 ASGMVAVS-----RAGDWFGSPVNVASRV-TGVARPGAVLVADSVREALGADPADGFGQ 353
DB 343 DTGSAASGLVGRSTLAYDMWGSADVAVYQVQSGSPQPG-IYVTSRVHEVMQETLD----- 396
QY 354 WSFAGPRRLRGIRGDRVRLFRVR 375
DB 397 --FVAAGEVVGGERGVETVWRRLQ 416

RESULT 13

AA14896

ID AA14896 standard; protein; 419 AA.

XX AC AA14896;

XX DT 25-OCT-1999 (first entry)

XX DE M. vaccae antigen GV-38B 3' amino acid sequence.

XX KW Mycobacterium vaccae protein; antigen; T cell activation; cytokine;
XX KW dendritic cell maturation; infectious disease; immune disorder; cancer;
KW respiratory system; mycobacterial infection; allergy; tuberculosis;
KW leprosy; sarcoidosis; lung cancer; asthma; skin disorder; psoriasis;
KW dermatitis; eczema; alopecia areata; skin cancer; basal carcinoma;
KW squamous cell carcinoma; melanoma.

XX OS Mycobacterium vaccae.

XX PN W09932634-A2.

XX PD 01-JUL-1999.

XX PF 23-DEC-1998; 98WO-NZ000189.

XX PR 23-DEC-1997; 97US-00996624.

XX PR 23-DEC-1997; 97US-00997080.

XX PR 23-DEC-1997; 97US-00997362.

PR 11-JUN-1998; 98US-00095855.
PR 17-SEP-1998; 98US-00156181.
PR 04-DEC-1998; 98US-00205426.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI Tan P, Watson J, Visser ES, Skinner MA, Prestidge RL;
XX WPI; 1999-430163/36.
DR N-PSDB; AAZ11349.
XX
PT Enhancing immune response to an antigen.
XX
PS Example 15; Page 197-198; 243pp; English.
XX
CC The invention provides heat-killed Mycobacterium vaccae, or recombinant
CC M. vaccae proteins. The M. vaccae proteins may be employed to activate T
CC cells and natural killer cells, to stimulate the production of cytokines,
CC and to enhance the expression of co-stimulatory molecules on dendritic cells
CC and monocytes, and to enhance dendritic cell maturation and function. The
CC proteins can be expressed by standard recombinant methodology.
CC Pharmaceutical compositions comprising the proteins or nucleic acid
CC sequences encoding the proteins can be used for the treatment
CC prevention, and detection of disorders including infectious diseases,
CC immune disorders and cancer. In particular, the compounds and methods are
CC used for treatment of diseases of the respiratory system, such as
CC mycobacterial infections, asthma, allergies, tuberculosis, leprosy,
CC sarcoidosis and lung cancers, and disorders of the skin such as
CC psoriasis, atopic dermatitis, eczema, allergic contact dermatitis, cell
CC alopecia areata, and skin cancers such as basal carcinoma, squamous cell
CC carcinoma and melanoma
XX
SQ Sequence 419 AA;

Query Match 7.6%; Score 151; DB 2; Length 419;

Best Local Similarity 23.3%; Pred. No. 1.6e-05;

Matches 103; Conservative 60; Mismatches 163; Indels 116; Gaps 19;

QY 14 DDLGLDGLGTA-----RAERAKLVLEWLLLEQGITPDEI-----RATNPP 51

DB 11 DTGMDGTETILVGPDLNMRSDSRFLFRENREKFLADVVEGG-TPEVADESVDRRGTTLV 69

QY 52 LLLATRLHV---GDDGT-----YVSAREISENYGVDLLELQVQRAVGLARVDDPAVV 102

DB 70 QPVTTSRVEEAQRNGTGTIEDDYLGEALQAYSPVDLPGLHWI-----VAKID----- 119

QY 103 HMRADGEAAARQRFVELGLNPDQVVLVVRVLAEGLSHAAEAAMRYTALRAIMRPGATELD 162

DB 120 ----TDEAFAPVAQFTR-----TLVLTSTVIIIFGVSLAAMLLARLFVVRPIRLQAGAO 169

QY 163 IAKGSOALVSQIVPL-----LGPMDQLFMQLRHMMETEAVNAGERAAKGKPL----- 210

DB 170 ISGGDYRLALPVLRSRDEFGLTAFNDM-----SRNLSIKDELLEGERAENQRLMSLMPE 225

QY 211 ----PGARQVTVAFADLVGFTOLGEVVSABEL-----GHLAGLAGLARDL 252

DB 226 PVMQRYLDEETIAQDHKNVTIFADMMGLDLSRLTSEELMVVNDLTQFDAAAESL 285

QY 253 TAPPWFYIKTIGDAMVLCPPAPLDTVLKLV-----EVVD-----TDNNFPRLRAGV 301

DB 286 ---GVDHVTFLHDGYLASCGLGVPRLDNVRRTVNFALENDRIIDRHAESGHDRLRAGI 342

QY 302 ASGMVAVS-----RAGDWFGSPVNVASRV-TGVARPGAVLVADSVREALGADPADGFGQ 353

DB 343 DTGSAASGLVGRSTLAYDMWGSADVAVYQVQSGSPQPG-IYVTSRVHEVMQETLD----- 396

QY 354 WSFAGPRRLRGIRGDRVRLFRVR 375

DB 397 --FVAAGEVVGGERGVETVWRRLQ 416

RESULT 14

ABB73502

ID ABB73502 standard; protein; 419 AA.
XX ABB73502;
AC ABB73502;
XX
DT 08-APR-2002 (first entry)
XX
DE M vaccae GV-388 partial protein SEQ ID NO: 125.
XX
KW Skin disorder; psoriasis; atopic dermatitis; allergic contact dermatitis;
KW alopecia areata; skin cancer; Mycobacterium vaccae; melanoma; cytostatic;
KW antipsoriatic; dermatological; antiinflammatory; antiallergic;
KW Th2 immune response; immunomodulatory.
XX
OS Mycobacterium vaccae.
XX
FN US6328978-B1.
XX
PD 11-DEC-2001.
XX
PF 02-JUN-1999; 99US-00324542.
XX
PR 23-DEC-1997; 97US-00997080.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI Watson JD, Tan PLJ, Prestidge R;
XX
DR WPI; 2002-138361/18.
DR N-PSDB; ABL36255.
XX
XX Inhibiting skin inflammation associated with skin disorder e.g.
PT psoriasis, by administering composition comprising delipidated and
PT delglycolipidated Mycobacterium vaccae cells or Mycobacterium vaccae
PT culture filtrate.
XX
PS Example 9; Col 139-142; 116pp; English.
XX
XX The present invention relates to a method of inhibiting skin inflammation
CC associated with a skin disorder selected from psoriasis, atopic
CC dermatitis and allergic contact dermatitis, which involves administering
CC a composition containing delipidated and delglycolipidated Mycobacterium
CC vaccae cells or M. vaccae culture filtrate. The skin disorder to be
CC treated may also include alopecia areata, and skin cancers such as basal
CC cell carcinoma, squamous cell carcinoma and melanoma. The composition
CC acts by inhibiting the Th2 immune response. The present sequence is a
CC protein described in the exemplification of the invention
XX
SQ Sequence 419 AA;
Query Match 7.6%; Score 151; DB 5; Length 419;
Best Local Similarity 23.3%; Pred.No.1.6e-05;
Matches 103; Conservative 60; Mismatches 163; Indels 116; Gaps 19;
QY 14 DDLGDLGGTA-----RAERAKLVWLEEQGITPDEI-----RATNPP 51
DB 11 DTGMDGTGETILVGPDLNMRSDSRLEFRNREKFLADVVEGG-TPPEVADESVDRRGTTLV 69
QY 52 LLLATRLV-----GDGT-----YVSAREISENYGVLDLELQVRQAVGLARVDDPDVAV 102
DB 70 QPVTTTSVEEAQRGNTGTTIEDDYLGHQALQAYSFVDPGLHWTI-----VAKID----- 119
QY 103 HMRADGAAARQAREVELGNPDVVLVVRVLAEGLSHAAEAMRYTAL EATMRPGATELD 162
DB 120 -----TDEAFAPVAQTR-----TLVLTSTVIFGVSLAAMLLARLFVRPIRRLOAQAO 169
QY 163 IAKGSOALVSQIVPL-----LGPMTQDMLFMQLRHMETEAVNAGERAAAGKPL----- 210
DB 170 ISGGDYRLALPVLSDREFGDLTTAFNDM-----SRNLSIKDELGGERAENQRLMLSLMPE 225
QY 211 -----PCARQVTVAFADLVGFTQLGVEVSAEEL-----GHLAAGLAGLARDL 252
DB 226 PWMQRYLGDGEETIAQDHKNVTVFADMMGLDELSEMLTSEELVMVVVNDLTRQFDAAAE 285

QY 253 TAPPVWFIKTIGDAVMLVCPDPAPLLDTVLKLV-----EVVD-----TDNNFRLRAGV 301
DB 286 ---GVDHVRTLHDGYLASCGLGVPRLDNVVRTVNFATEMDRIIDRHAAESGHDRLRLRAGI 342
QY 302 ASGMAYS-----RAGDWFQSPVNVASRV-TGVARPGAVLVADSVREAI.GDAI.EADGFG 353
DB 343 DTGSAASGLVGRSTLAYUMGSAVDVAVQVQRGSPQPG-IYVTSRVHEVMQETLD----- 396
QY 354 WSPAGPRLRGIRGDVRLFRVR 375
DB 397 --FVAAGEVVGGERGVETVWRLQ 416
RESULT 15
ADJ83100
ID ADJ83100 standard; protein; 194 AA.
XX
AC ADJ83100;
XX
DT 06-MAY-2004 (first entry)
XX
DE Adenylyl and guanylyl cyclase catalytic domain protein - SEQ ID 91.
XX
XX NOVX; cytostatic; antiarteriosclerotic; antidiabetic; antiasthmatic;
KW antiinflammatory; respiratory; antiarthritic;
KW dermatological; antibacterial; cerebroprotective; vasotropic; cardiac;
KW haemostatic; hypotensive; hepatotropic; neuroprotective; anorectic;
KW nootropic; antiulcer; muscular; immunosuppressive; gynaecological;
KW antiparkinsonian; endocrine; ophthalmological; osteopathic;
KW antiparkinsonian; endocrine; ophthalmological; osteopathic;
KW cell signal processing; metabolic pathway; asthma; allergy; emphysema;
KW autoimmune; graft-versus-host; arthritis; cancer; stroke; haemophilia;
KW obesity; Alzheimer's; pain; chromosome mapping; tissue typing;
KW adenylyl cyclase; enzyme; guanylyl cyclase; catalytic domain.
XX
OS Unidentified.
XX
PN US2003170630-A1.
XX
PD 11-SEP-2003.
XX
PF 21-DEC-2001; 2001US-00032189.
XX
PR 21-DEC-2000; 2000US-0257495P.
PR 22-DEC-2000; 2000US-0258171P.
PR 20-FEB-2001; 2001US-0269940P.
PR 08-MAR-2001; 2001US-0274192P.
PR 22-MAR-2001; 2001US-0277826P.
PR 29-MAR-2001; 2001US-0279840P.
PR 11-APR-2001; 2001US-0282981P.
PR 13-APR-2001; 2001US-0283656P.
PR 31-JUL-2001; 2001US-0309247P.
PR 10-AUG-2001; 2001US-0311754P.
PR 17-AUG-2001; 2001US-0313331P.
XX
XX (ALSO/) ALSOROOK J P.
PA (TCHU/) TCHERNEV V T.
PA (LIUX/) LIU X.
PA (SPYT/) SPYTEK K A.
PA (ZERH/) ZERHUSEN B D.
PA (PATT/) PATTURAJAN M.
PA (LEPL/) LEPLEY D M.
PA (BURG/) BURGESS C E.
PA (SHIM/) SHIMKETS R A.
PA (GROS/) GROSSE W M.
PA (SZEK/) SZEKERES E S.
PA (VERN/) VERNET C A M.
PA (LILL/) LI L.
PA (CASM/) CASMAN S J.
PA (BOLD/) BOLDOG F L.
PA (GORM/) GORMAN L.
PA (GANG/) GANGOLLI E A.

PA (PERN/) FERNANDES E R.
 PA (RIEG/) RIEGER D K.
 PA (EDIN/) EDINGER S R.
 PA (GUNT/) GUNTHER E.
 PA (MILL/) MILLET I.
 PA (SCIO/) SCIORE P.
 PA (ELLE/) ELLERMAN K.
 PA (MACD/) MACDOUGALL J R.
 PA (SMIT/) SMITHSON G.
 XX
 PI Alsobrook JP, Tcherev VT, Liu X, Spytek KA, Zerhusen BD:
 PI Patturajan M, Lepley DM, Burgess CE, Shinkets RA, Grosse WM;
 PI Szekeres ES, Vernhet CAM, Li L, Casman SJ, Boldog FL, Gorman L;
 PI Gangolli EA, Fernandes ER, Rieger DK, Edinger SR, Gunther E;
 PI Millet I, Sciore P, Ellerman K, Macdougall JR, Smithson G;
 XX
 DR WPI; 2003-898249/82.
 XX
 PT New NOVX polypeptides and nucleic acid molecules, useful for diagnosing,
 PT preventing or treating NOVX-associated polypeptide disorder, e.g.
 PT cardiomyopathy, atherosclerosis, diabetes, cancer, Parkinson's disease or
 PT asthma.
 XX
 PS Disclosure; SEQ ID NO 91; 263pp; English.
 XX
 CC The invention relates to a novel isolated NOVX polypeptide. The
 CC polypeptide demonstrates cytostatic, antiarteriosclerotic, antidiabetic,
 CC antiasthmatic, anti-allergic, anti-inflammatory, respiratory,
 CC antiarthritic, dermatological, antibacterial, cerebroprotective,
 CC vasotropic, cardiant, haemostatic, hypotensive, hepatotropic,
 CC neuroprotective, anorectic, nootropic, antiulcer, muscular,
 CC immunosuppressive, gynaecological, antipsoriatic, endocrine,
 CC ophthalmological, osteopathic, antiparkinsonian, anticonvulsant,
 CC tranquiliser, analgesic, nephrotropic, antiinfertility and antilipaemic
 CC activities. The NOVX polypeptide, nucleic acid or antibody of the
 CC invention may be useful for treating or preventing a NOVX-associated
 CC disorder, such as cardiomyopathy, atherosclerosis, diabetes or a disorder
 CC related to cell signal processing and metabolic pathway modulation.
 CC Furthermore, the NOVX polypeptides may be useful for diagnosing, treating
 CC or preventing diseases such as asthma, allergies, emphysema, autoimmune
 CC disease, graft-versus-host disease, arthritis, cancer, stroke,
 CC haemophilia, obesity, Alzheimer's disease and pain. The nucleic acids may
 CC be used as hybridisation probes, in chromosome mapping, tissue typing,
 CC preventive medicine or pharmacogenomics. The current sequence is that of
 CC a protein of the invention which is related to human NOVX protein.
 XX
 SQ Sequence 194 AA;

Query Match 7.4%; Score 148.5; DB 7; Length 194;
 Best Local Similarity 31.2%; Pred. No. 9.3e-06;
 Matches 58; Conservative 23; Mismatches 72; Indels 33; Gaps 8;

QY 184 DMLFMQLRHMMETEVNAGERAGKPLPGAR--QVTVAFADLVGFTQLGEVVSARELGHIL 241
 DB 7 DRLLDQLLPASVAESLKR-----GEVPAPSYDEVITFLTIDVGFALSAAATPEQVVTLL 62
 QY 242 AGRL-AGLARDLTAPPVWFITKIGDAMLVCPDPA-----PLLDTV--LKLVE 286
 DB 63 LNDLYSRFDRIIDRHGGYKVKITIGDAYVMVSGLPTAALVQHAEALADNVESLKTVL 122
 QY 287 VVDTDNFPRLRAGVASGMASVSRAG-----DMFGSPVNVASRVTVGVRPGAVLVAD 337
 DB 123 VQHRGNGL-RVRIGIHTGPVW--AGVVGITMPRYCLFGDTVNLASRMESVGDPCQIQVSE 179
 QY 338 SVREAL 343
 DB 180 ETYSLL 185

Search completed: October 15, 2005, 04:26:23
 Job time : 107.54 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 01:42:44 ; Search time 32.5955 Seconds
(without alignments)
909.197 Million cell updates/sec

Title: US-10-617-038-7
Perfect score: 1999
Sequence: 1 VTDHVREADDANIDLLGLD.....ATRTGGGAQDDDLGSSP 397

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A-COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B-COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCITUS-COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	157.5	7.9	486	US-09-902-540-12750	Sequence 12750, A
2	157	7.9	670	US-08-997-080-178	Sequence 178, App
3	157	7.9	670	US-08-997-362-178	Sequence 178, App
4	157	7.9	670	US-09-095-855-178	Sequence 178, App
5	157	7.9	670	US-09-324-542-178	Sequence 178, App
6	157	7.9	670	US-09-205-426-178	Sequence 178, App
7	157	7.9	722	US-08-997-080-174	Sequence 174, App
8	157	7.9	722	US-08-997-362-174	Sequence 174, App
9	157	7.9	722	US-09-095-855-174	Sequence 174, App
10	157	7.9	722	US-09-324-542-174	Sequence 174, App
11	157	7.9	722	US-09-205-426-174	Sequence 174, App
12	151	7.6	419	US-08-997-080-125	Sequence 125, App
13	151	7.6	419	US-08-997-362-125	Sequence 125, App
14	151	7.6	419	US-09-095-855-125	Sequence 125, App
15	151	7.6	419	US-09-324-542-125	Sequence 125, App
16	139.5	7.0	645	US-09-205-426-125	Sequence 125, App
17	139.5	7.0	645	US-09-949-016-9562	Sequence 9562, App
18	130.5	6.5	535	US-09-252-991A-25604	Sequence 25604, A
19	128	6.4	1144	US-09-949-016-6446	Sequence 6446, App
20	128	6.4	1149	US-09-949-016-7682	Sequence 7682, App
21	127	6.4	450	US-09-902-540-16316	Sequence 16316, App
22	127	6.4	823	US-09-252-991A-31026	Sequence 31026, A
23	127	6.4	1047	US-09-949-016-9830	Sequence 9830, App
24	127	6.4	1047	US-09-949-016-9831	Sequence 9831, App
25	126.5	6.3	543	US-09-252-991A-28875	Sequence 28875, A
26	126.5	6.3	1304	US-09-489-039A-13449	Sequence 13449, A
27	126	6.3	422	US-09-902-540-16169	Sequence 16169, A

ALIGNMENTS

RESULT 1

US-09-902-540-12750
; Sequence 12750, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12750
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12750

Query Match 7.9%; Score 157.5; DB 4; Length 486;

Best Local Similarity 27.4%; Pred. No. 1.1e-07;

Matches 93; Conservative 44; Mismatches 124; Indels 79; Gaps 17;

QY 80 LELLQVRQVAVGLARVDDPDVAVVHMRADGEAAARQRFVELGLNPDQVVLVVRVLAEGUS 139

DB 178 LESMGRLOPVIVLSDEMPG-----LDG-----VEL-----CRRLEKEGPS 212

QY 140 HAAEAMRYTA--LEAI-----MRPGATELDIAKGSQ-----ALVSOIVPLGLPMIQDMLFMQ 189

DB 213 HAVPAIIITANDVEAVERVRGRAGADDYVNSASWDELAARIESVVRRTGE--TERMRKLF 271

QY 190 LRHMMETEAVNAGERAAAGKPLPG-ARQVTVAFDLVGFTQLGEVVSABELHLAGRLAGL 248

DB 272 ARYTSADVVEELTKSADAVLTKGEKEVTLTFADIRNFTGLAESLPPEQVGVNLQVLGR 331

QY 249 ARD--LTAPPVWFIKTIGDAVNLVCPDPAPLDTVLKIVE-----VVD--TDNNFP----- 295

DB 332 MSDAVLTCCGT-LDKFLGDLGMAVFGAPVAPRDDALRALQSAKMMMDALTDLRIEAEAW 390

QY 296 -----RLRAGVASGNMVSAGD-----WFGSPVNVASRVTCVAPPGAV 333

DB 391 AAGGREGQPLVELGLGINSVGVV--AGNIGSTVRAEYTCIGDAVNVAAARLALCAGPGEI 448

QY 334 LVADSVRALGDAPADGFGFMSFAGPRRLRGIRGOVRLFR 373

DB 419 LVGETRE-LVDANET---AFEDLPPVRLKGQKQVPVPLR 484

28 126 6.3 882 4 US-09-252-991A-17653 Sequence 17653, A
29 122 6.1 408 4 US-09-252-991A-31315 Sequence 31315, A
30 121.5 6.1 2216 4 US-09-902-540-12221 Sequence 12221, A
31 121 6.1 4551 3 US-09-320-878-1 Sequence 1, Appli
32 121 6.1 4551 4 US-09-141-908-2 Sequence 2, Appli
33 121 6.1 4551 4 US-09-657-440-1 Sequence 1, Appli
34 121 6.1 4613 3 US-09-105-537-31 Sequence 31, Appli
35 121 6.1 11877 3 US-09-105-537-6 Sequence 6, Appli
36 120.5 6.0 640 4 US-09-252-991A-27604 Sequence 27604, A
37 120.5 6.0 6095 3 US-09-144-085-2 Sequence 2, Appli
38 120 6.0 1114 4 US-09-252-991A-24965 Sequence 24965, A
39 120 6.0 2214 4 US-09-902-540-15988 Sequence 15988, A
40 119.5 6.0 805 4 US-09-252-991A-32752 Sequence 32752, A
41 119.5 6.0 1253 3 US-08-864-785-2 Sequence 2, Appli
42 117 5.9 683 4 US-09-252-991A-32144 Sequence 32144, A
43 115.5 5.8 650 4 US-09-252-991A-32334 Sequence 32334, A
44 115.5 5.8 1144 3 US-08-726-214-6 Sequence 6, Appli
45 115.5 5.8 1248 3 US-08-726-214-16 Sequence 16, Appli

```
RESULT 2
US-08-997-080-178
; Sequence 178, Application US/08997080
; Patent No. 5968524
; GENERAL INFORMATION:
; APPLICANT: WATSON, JAMES D.
; APPLICANT: TAN, PAUL L.J.
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,080
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 178:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 670 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-997-080-178

Query Match          7.9%; Score 157; DB 2; Length 670;
Best Local Similarity 23.9%; Pred. No. 2.le-07;
Matches 105; Conservative 55; Mismatches 159; Indels 120; Gaps 20;

Qy 14 DDLGLDGGTA-----RAERAKLVEMWLEQGITPDEI-----RATNPP 51
Db 261 DTGMDGTGETILVGPDLNMRSDSRFLFRENREKFLADVVEGG-TPEVADESVDRRGTTLV 319

Qy 52 LLLATRLHV-----GDDGT-----YVSAREISENYGVDELLELQVRQVAVGLARVDDPDVAV 102
Db 320 QPVTTSRVEEAQRGTGTTIEDDYLGHQALQAYSVDLPLGLHWI-----VAKID----- 369

Qy 103 HMRADGAAAAAQRQVFLVGLNPDQVVLVVRVLAEGLSHAAMRYTALEA1MRPGATELD 162
Db 370 ----TDEAFAPVAQFTR-----TLVLTSTV11IFGVSLAAMLLARLFVRPIRRLQAGAAQ 419

Qy 163 IAKGSOALVSQIVPL-----LGPMTQDMLFQOLRHHMETEAVNAGERAAGKPL----- 210
Db 420 ISGGDYRLALPVLRSDFEGDITTAENDM-----SRNLSIKDELGEERAENQRLMSLMP 475

Qy 211 -----PGARQVTVAFADLVGFTQLGGEWSAEEL-----GHLAGRLAGLARDL 252
Db 476 PVMQRYLDGEETIAQDHKNVTVIPADMMGLDLSRLMTSEELVVVNDLTLQFQAAAAESL 535

Qy 253 TAPPVWFKTIGDAVMLVCPDPAPLDITVLKLV-----EVDV-----TDNNFPLRAGV 301
Db 536 ---GVDHVRTLHDGYLASCGLGVPRLDNVRRTVNFPAIEMDRIIDRHAABESGHDLRLRAGI 592
```

```
Qy 302 ASGMVAVS-----RAGDMFGSPVNVASRV-TGVARPGAVLVADSVRE-----ALG 344
Db 593 DTGSAASGLVGRSTLAYDMGSAVDVANQVORGSPQPG-IYVTSRVHEVMOETLDFVAA 651

Qy 345 DAPEADGQFQ--WSPAGPRR 361
Db 652 EVVGERGVETVWR1LQGHRR 670

RESULT 3
US-08-997-362-178
; Sequence 178, Application US/08997362
; Patent No. 5985287
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Scott, Linda
; APPLICANT: Prestidge, Rosa
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,362
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
; FILING DATE: June 12, 1997
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
; FILING DATE: August 29, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 178:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 670 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-997-362-178

Query Match          7.9%; Score 157; DB 2; Length 670;
Best Local Similarity 23.9%; Pred. No. 2.le-07;
Matches 105; Conservative 55; Mismatches 159; Indels 120; Gaps 20;

Qy 14 DDLGLDGGTA-----RAERAKLVEMWLEQGITPDEI-----RATNPP 51
Db 261 DTGMDGTGETILVGPDLNMRSDSRFLFRENREKFLADVVEGG-TPEVADESVDRRGTTLV 319

Qy 52 LLLATRLHV-----GDDGT-----YVSAREISENYGVDELLELQVRQVAVGLARVDDPDVAV 102
Db 320 QPVTTSRVEEAQRGTGTTIEDDYLGHQALQAYSVDLPLGLHWI-----VAKID----- 369

Qy 103 HMRADGAAAAAQRQVFLVGLNPDQVVLVVRVLAEGLSHAAMRYTALEA1MRPGATELD 162
```


QY 103 HMRADGEAARQRFVELGNPDQVVLVVRVLAEGLSHAABAMRYTALAEIMRPGATELD 162
DB 370 ----TDEAFAPVAQTR-----TLVLSVIIIFGVSLAAMLLARLFRPIRRLQAQAQ 419
QY 163 IAKGSOALVSQIVPL-----LGPMTQDMLFMQLRHHMETEAVNAGERAAAGKPL----- 210
DB 420 ISGGDYRLALPVLRSDFEGLTTAFNDM-----SRNLSIKDELGGERAENQRLMSLME 475
QY 211 -----PGARQVTVAFADLVGFTQLGEVVSABEL-----GHLAAGRLAGLARDL 252
DB 476 PVMQRYLDCBETIAQDHKNVTIVFADMMGLDELSSRLTSEELMVVVNDLTRQFDDAAESL 535
QY 253 TAPPVWFIKTIGDAMLVCPDPAPLLDTVLKLV-----EVVD-----TDNNFRLRAGV 301
DB 536 ---GVDHVRTLHDGYLASCGLVGPRLDNVRRVTFNFAIEMDRIIDRHAAESGHDRLRLRAGI 592
QY 302 ASGMAYS-----RAGDWFQSPVNVASRV-TGVARPGAVLVADSVRE-----ALG 344
DB 593 DTGSAASGLVGRSTLAYDMWGSADVAVNQVQRGSPQPG-IYVTSRVHEVMQETLDFVAAG 651
QY 345 DAPADGFGQ--WSFAGPRR 361
DB 652 EVVGERGVETVNRLOQHRR 670

RESULT 6

US-09-205-426-178
; Sequence 178, Application US/09205426
; Patent No. 6406704
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; TITLE OF INVENTION: Compounds and Methods for Treatment and
; FILE OF INVENTION: Diagnosis of Mycobacterial Infections
; FILE REFERENCE: 11000.1002c4
; CURRENT APPLICATION NUMBER: US/09/205,426
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: 09/095,855
; EARLIER FILING DATE: 1998-06-11
; EARLIER APPLICATION NUMBER: 08/997,362
; EARLIER FILING DATE: 1997-12-23
; EARLIER APPLICATION NUMBER: 08/873,970
; EARLIER FILING DATE: 1997-06-12
; EARLIER APPLICATION NUMBER: 08/705,347
; EARLIER FILING DATE: 1996-08-29
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 178
; LENGTH: 670
; TYPE: PR1
; ORGANISM: Mycobacterium vaccae
US-09-205-426-178

Query Match 7.9%; Score 157; DB 4; Length 670;
Best Local Similarity 23.9%; Pred. No. 2.1e-07;
Matches 105; Conservative 55; Mismatches 15; Indels 120; Gaps 20;

QY 14 DLLGLDGLGGA-----RAERAKLVWELLEGITPDEI-----RATNPP 51
DB 261 DTGMGDTGETILVGPDLNMRSDSRFLFRENREKFLADVVEGG-TPEVADESVDRRGTTLV 319
QY 52 LLLATRLHV-----GDDGT-----YVSAREISENYGVLDLELLQRYORAVGLARVDDPDAV 102
DB 320 QPVVTRTSVEEAQRGNTGTTIEDYLGHEALQAYSFVDPGLGHWI-----VAKID----- 369
QY 103 HMRADGEAARQRFVELGNPDQVVLVVRVLAEGLSHAABAMRYTALAEIMRPGATELD 162
DB 370 ----TDEAFAPVAQTR-----TLVLSVIIIFGVSLAAMLLARLFRPIRRLQAQAQ 419
QY 163 IAKGSOALVSQIVPL-----LGPMTQDMLFMQLRHHMETEAVNAGERAAAGKPL----- 210
DB 420 ISGGDYRLALPVLRSDFEGLTTAFNDM-----SRNLSIKDELGGERAENQRLMSLME 475

QY 211 -----PGARQVTVAFADLVGFTQLGEVVSABEL-----GHLAAGRLAGLARDL 252
DB 476 PVMQRYLDCBETIAQDHKNVTIVFADMMGLDELSSRLTSEELMVVVNDLTRQFDDAAESL 535
QY 253 TAPPVWFIKTIGDAMLVCPDPAPLLDTVLKLV-----EVVD-----TDNNFRLRAGV 301
DB 536 ---GVDHVRTLHDGYLASCGLVGPRLDNVRRVTFNFAIEMDRIIDRHAAESGHDRLRLRAGI 592
QY 302 ASGMAYS-----RAGDWFQSPVNVASRV-TGVARPGAVLVADSVRE-----ALG 344
DB 593 DTGSAASGLVGRSTLAYDMWGSADVAVNQVQRGSPQPG-IYVTSRVHEVMQETLDFVAAG 651
QY 345 DAPADGFGQ--WSFAGPRR 361
DB 652 EVVGERGVETVNRLOQHRR 670

RESULT 7

US-08-997-080-174
; Sequence 174, Application US/08997080
; Patent No. 5968524
; GENERAL INFORMATION:
; APPLICANT: WATSON, JAMES D.
; APPLICANT: TAN, PAUL L. J.
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESS: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,080
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 174:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 722 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-997-080-174

Query Match 7.9%; Score 157; DB 2; Length 722;
Best Local Similarity 23.9%; Pred. No. 2.4e-07;
Matches 105; Conservative 55; Mismatches 15; Indels 120; Gaps 20;

QY 14 DLLGLDGLGGA-----RAERAKLVWELLEGITPDEI-----RATNPP 51
DB 313 DTGMGDTGETILVGPDLNMRSDSRFLFRENREKFLADVVEGG-TPEVADESVDRRGTTLV 371
QY 52 LLLATRLHV-----GDDGT-----YVSAREISENYGVLDLELLQRYORAVGLARVDDPDAV 102
DB 372 QPVVTRTSVEEAQRGNTGTTIEDYLGHEALQAYSFVDPGLGHWI-----VAKID----- 421

```
QY 103 HMRADGAAARQRFVELGLNPDQVVLVVRVLAEGLSHAAEAMRYTALEA1MRPGATELD 162
Db 422 ----TDEAFAPVAQFTR-----TLVLSVTIIIFGVSLAAMLRLFRVPIRRLQAGAQ 471
QY 163 IAKGSOALVSQIVPL-----LGPMDQMLFMQLRHMMETEA1VNAAGRAAGKPL----- 210
Db 472 ISGGDYRLALPVLRSDEFGDLTTAFNDM-----SRNLSIKDELLGGERAENQRLMLSLMPE 527
QY 211 -----PGARQVTVAFADLVGFTOLGEVWSAEEL-----GHLA9RLAGLARDL 252
Db 528 PVMQRYLDGEETIAQDHKNVTVIIFADMMGLDLSRMLTSEELMVVVNDLTRQFDDAAEESL 587
QY 253 TAPPVWF1KTIGDAVMLVCPDPAPLLDTVLKLV-----EVVD-----TDNNFPRLRAGV 301
Db 588 ---GVDHVRTLHDGVLASCGLVGPRLDNVRRTVNF1EMDRIIDHAAESGHDLRLRAGI 644
QY 302 ASGMVAVS-----RAGDWFSGPVNVASRV-TGVARPGAVLVAADVRE-----ALG 344
Db 645 DTGSAASGLVGRSTLAYDMGSAVDVANQVQSGSPQPG-IYVTSRVHEVMQETLDFVAAG 703
QY 345 DAPEADGFGQ--WSFAGPRR 361
Db 704 EVVGERGVETVWR1LQGHRR 722

RESULT 8
US-08-997-362-174
; Sequence 174, Application US/08997362
; Patent No. 5985287
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Scott, Linda
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
; TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,362
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
; FILING DATE: June 12, 1997
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
; FILING DATE: August 29, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 174:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 722 amino acids
; TYPE: amino acid
```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-997-362-174

Query Match 7.9%; Score 157; DB 2; Length 722;
Best Local Similarity 23.9%; Pred. No. 2.4e-07;
Matches 105; Conservative 55; Mismatches 159; Indels 120; Gaps 20;

QY 14 DLLIGDLGGTA-----RAERAKLVLEWLEEQGITPDEI-----RATNPP 51
Db 313 DTGMDGTGETILVGPDNLMRSDSR1FRENREKFLADVVEGG-TPEVADESVDREGTTLV 371
QY 52 LLLATRLHV-----GDDGT-----YVSARBEISENYVDLELLQVORAVGLARVDDPDVV 102
Db 372 QPVTTSRVEEAQRGTGTTIBDDYLGHLEALQAYSPVDLPGLHWI-----VAKID----- 421
QY 103 HMRADGAAARQRFVELGLNPDQVVLVVRVLAEGLSHAAEAMRYTALEA1MRPGATELD 162
Db 422 ----TDEAFAPVAQFTR-----TLVLSVTIIIFGVSLAAMLRLFRVPIRRLQAGAQ 471
QY 163 IAKGSOALVSQIVPL-----LGPMDQMLFMQLRHMMETEA1VNAAGRAAGKPL----- 210
Db 472 ISGGDYRLALPVLRSDEFGDLTTAFNDM-----SRNLSIKDELLGGERAENQRLMLSLMPE 527
QY 211 -----PGARQVTVAFADLVGFTOLGEVWSAEEL-----GHLA9RLAGLARDL 252
Db 528 PVMQRYLDGEETIAQDHKNVTVIIFADMMGLDLSRMLTSEELMVVVNDLTRQFDDAAEESL 587
QY 253 TAPPVWF1KTIGDAVMLVCPDPAPLLDTVLKLV-----EVVD-----TDNNFPRLRAGV 301
Db 588 ---GVDHVRTLHDGVLASCGLVGPRLDNVRRTVNF1EMDRIIDHAAESGHDLRLRAGI 644
QY 302 ASGMVAVS-----RAGDWFSGPVNVASRV-TGVARPGAVLVAADVRE-----ALG 344
Db 645 DTGSAASGLVGRSTLAYDMGSAVDVANQVQSGSPQPG-IYVTSRVHEVMQETLDFVAAG 703
QY 345 DAPEADGFGQ--WSFAGPRR 361
Db 704 EVVGERGVETVWR1LQGHRR 722

RESULT 9
US-09-095-855-174
; Sequence 174, Application US/09095855
; Patent No. 6160093
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Compounds and Methods for
; TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 208
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,855
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; APPLICATION NUMBER: 08/873,970
```

```
; FILING DATE: 12-JUN-1997
; APPLICATION NUMBER: 08/997,362
; FILING DATE: 23-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 174:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 722 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-095-855-174

Query Match          7.9%; Score 157; DB 3; Length 722;
Best Local Similarity 23.9%; Pred. No. 2.4e-07;
Matches 105; Conservative 55; Mismatches 159; Indels 120; Gaps 20;

Qy 14 DDLGLDGLGTA-----RAERAKLVEMWLEQGITPDEI-----RATNPP 51
Db 313 DTGMGDTGETILVCPDNLMSRSLFRENREKFLADVVEGG-TPEVADESVDRRGTTLV 371
Qy 52 LLLATRLHV-----GDDGT-----YVSAREISENYGVDELLELQVORAVGLARVDDPDVAV 102
Db 372 QPVTTRSVEEAQRGNTGTTIEDDYLGHEALQAYSVDLPGLHWI-----VAKID----- 421
Qy 103 HMRADGEAARAQRFEVLGNPDQVVLVVRVLAESGLSHAAEAMRYTALEAIRMRTGATELD 162
Db 422 ----TDEAFAPVAQFTR-----TLVLSTVIIIFGVSLAAMLLARLFVRPIRRLQAGAQQ 471
Qy 163 IAKGSQALVSQIVPL-----LGPMTQDMLFMQLRHHMETEAVNAGERAAGKPL----- 210
Db 472 ISGGDYRLALPVLRSDEFGLDTTAFNDM-----SRNLSIKDELLGEERAENQRLMSLME 527
Qy 211 -----PGARQVTVAFADLVGFTQLGEVVSABEEL-----CHLAGRLAGLARDL 252
Db 528 PVMQRYLDGEBETIAQDHKNVTVI FADMMGLDELSRMLTSEELMVVNDLTRQFDAAAESL 587
Qy 253 TAPPVWFKTIGDAVMLVCPDAPLDDTVLKL-----EVVD-----TDNNFPRLRAGV 301
Db 588 ---GVDHVRTLHDGYLASCGLGVPRLDNVRTVFAIEMDRIIDRHAESGHDRLRAGI 644
Qy 302 ASGNMVS-----RAGDMFGSPVNVASRV-TGVARPGAVLVADSVRE-----ALG 344
Db 645 DTGSAASGLVGRSTLAYDMWGSADVNAVQVORGSPQPG-IVVTSRVHEVMQETLDFVAAG 703
Qy 345 DAPADGFGQ--WSFAGPRR 361
Db 704 EVVGERGVETVMRLQGHRR 722

RESULT 10
US-09-324-542-174
; Sequence 174, Application US/09324542
; Patent No. 6328978
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L.J.
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: Of Immunologically-Mediated Skin Disorders
; FILE REFERENCE: 11000.1007c1
; CURRENT APPLICATION NUMBER: US/09/324,542
; CURRENT FILING DATE: 1999-06-02
; EARLIER APPLICATION NUMBER: US 08/997,080
; EARLIER FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FastSEQ for Windows Version 3.0
;
```

```
; SEQ ID NO 174
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
US-09-324-542-174

Query Match          7.9%; Score 157; DB 3; Length 722;
Best Local Similarity 23.9%; Pred. No. 2.4e-07;
Matches 105; Conservative 55; Mismatches 159; Indels 120; Gaps 20;

Qy 14 DDLGLDGLGTA-----RAERAKLVEMWLEQGITPDEI-----RATNPP 51
Db 313 DTGMGDTGETILVCPDNLMSRSLFRENREKFLADVVEGG-TPEVADESVDRRGTTLV 371
Qy 52 LLLATRLHV-----GDDGT-----YVSAREISENYGVDELLELQVORAVGLARVDDPDVAV 102
Db 372 QPVTTRSVEEAQRGNTGTTIEDDYLGHEALQAYSVDLPGLHWI-----VAKID----- 421
Qy 103 HMRADGEAARAQRFEVLGNPDQVVLVVRVLAESGLSHAAEAMRYTALEAIRMRTGATELD 162
Db 422 ----TDEAFAPVAQFTR-----TLVLSTVIIIFGVSLAAMLLARLFVRPIRRLQAGAQQ 471
Qy 163 IAKGSQALVSQIVPL-----LGPMTQDMLFMQLRHHMETEAVNAGERAAGKPL----- 210
Db 472 ISGGDYRLALPVLRSDEFGLDTTAFNDM-----SRNLSIKDELLGEERAENQRLMSLME 527
Qy 211 -----PGARQVTVAFADLVGFTQLGEVVSABEEL-----CHLAGRLAGLARDL 252
Db 528 PVMQRYLDGEBETIAQDHKNVTVI FADMMGLDELSRMLTSEELMVVNDLTRQFDAAAESL 587
Qy 253 TAPPVWFKTIGDAVMLVCPDAPLDDTVLKL-----EVVD-----TDNNFPRLRAGV 301
Db 588 ---GVDHVRTLHDGYLASCGLGVPRLDNVRTVFAIEMDRIIDRHAESGHDRLRAGI 644
Qy 302 ASGNMVS-----RAGDMFGSPVNVASRV-TGVARPGAVLVADSVRE-----ALG 344
Db 645 DTGSAASGLVGRSTLAYDMWGSADVNAVQVORGSPQPG-IVVTSRVHEVMQETLDFVAAG 703
Qy 345 DAPADGFGQ--WSFAGPRR 361
Db 704 EVVGERGVETVMRLQGHRR 722

RESULT 11
US-09-205-426-174
; Sequence 174, Application US/09205426
; Patent No. 6406704
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; TITLE OF INVENTION: Compounds and Methods for Treatment and
; TITLE OF INVENTION: Diagnosis of Mycobacterial Infections
; FILE REFERENCE: 11000.1002c4
; CURRENT APPLICATION NUMBER: US/09/205,426
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: 09/095,855
; EARLIER FILING DATE: 1998-06-11
; EARLIER APPLICATION NUMBER: 08/997,362
; EARLIER FILING DATE: 1997-12-23
; EARLIER APPLICATION NUMBER: 08/873,970
; EARLIER FILING DATE: 1997-06-12
; EARLIER APPLICATION NUMBER: 08/705,347
; EARLIER FILING DATE: 1996-08-29
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 174
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
US-09-205-426-174

Query Match          7.9%; Score 157; DB 4; Length 722;
Best Local Similarity 23.9%; Pred. No. 2.4e-07;
```



```
;
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
; FILING DATE: June 12, 1997
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
; FILING DATE: August 29, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 125:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-997-362-125

Query Match 7.6%; Score 151; DB 2; Length 419;
Best Local Similarity 23.3%; Pred. No. 4.3e-07;
Matches 103; Conservative 60; Mismatches 163; Indels 116; Gaps 19;

Qy 14 DLLGLDGLGGA-----RAERAKLVEMWLEQGITPDEI-----RATNPP 51
Db 11 DTGMGDTGETILVGPDLNMRSDSRFLFRENREKFLADVVEGG-TTPEVADESVDRRGTTLV 69
Qy 52 LLLATRLHV-----GDDGT-----YVSAREISENYGVDELLQVRQVAVGLARVDDPDVAV 102
Db 70 QPVVTRSVEEAQRGNTGTTIEDDYLGHQAQYSPVDLPGLHWI-----VAKID----- 119
Qy 103 HMRADGSAARQRFVBLGLNPDQVVLVVRVLAEGLSHAABMYRTALEAIMRPGATELD 162
Db 120 ----TDEAFAPVAQFTR-----TLVLTSTVIIIFGVSLAAMLLARLFVRPIRRLQAQAQ 169
Qy 163 IAKGSQALVSQIVPL-----LGPMTQDMLFMQLRHHMTEAVNAGERAAGKPL----- 210
Db 170 ISGGDYRLALPVLRSDFEGDLTTAFNDM-----SRNLSIKDELLGEERAENQRLMSLMPE 225
Qy 211 -----PGARQVTVAFADLVGFTQLGEVVSABEL-----EYVD-----TNNFPRLRAGV 301
Db 226 PVMQRYLDBEETIAQDHKNVTVIFADMMGLDELSEMLTSEELMVVVNDLTQFDAAAE 285
Qy 253 TAPPVWFIKTIGDAMLVCPDPAPLLDTVLKLV-----EYVD-----TNNFPRLRAGV 301
Db 286 ---GVDHVRTLHDGYLASCGLGVPRLDNVRVTNFAIEMDRIIDRHAESGHDLRLRAGI 342
Qy 302 ASGMVAVS-----RAGDWFSGPVNVASRV-TGVARPGAVLVADSVREALGDAPADG 353
Db 343 DTGSAAGSLVGRSTLAYDMMGSAVDVAVQVORGSPQPG-IYVTSRVHEVMQETILD----- 396
Qy 354 WSFAGPRRLRGIRGDVRLFRVR 375
Db 397 --FVAAGEVVGGRVETVWRLQ 416

RESULT 14
US-09-095-855-125
; Sequence 125, Application US/09095855
; Patent No. 6160093
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Compounds and Methods for
; TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 208
; CORRESPONDENCE ADDRESS:
```

```
;
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,855
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; APPLICATION NUMBER: 08/873,970
; FILING DATE: 12-JUN-1997
; APPLICATION NUMBER: 08/997,362
; FILING DATE: 23-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 125:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-095-855-125

Query Match 7.6%; Score 151; DB 3; Length 419;
Best Local Similarity 23.3%; Pred. No. 4.3e-07;
Matches 103; Conservative 60; Mismatches 163; Indels 116; Gaps 19;

Qy 14 DLLGLDGLGGA-----RAERAKLVEMWLEQGITPDEI-----RATNPP 51
Db 11 DTGMGDTGETILVGPDLNMRSDSRFLFRENREKFLADVVEGG-TTPEVADESVDRRGTTLV 69
Qy 52 LLLATRLHV-----GDDGT-----YVSAREISENYGVDELLQVRQVAVGLARVDDPDVAV 102
Db 70 QPVVTRSVEEAQRGNTGTTIEDDYLGHQAQYSPVDLPGLHWI-----VAKID----- 119
Qy 103 HMRADGSAARQRFVBLGLNPDQVVLVVRVLAEGLSHAABMYRTALEAIMRPGATELD 162
Db 120 ----TDEAFAPVAQFTR-----TLVLTSTVIIIFGVSLAAMLLARLFVRPIRRLQAQAQ 169
Qy 163 IAKGSQALVSQIVPL-----LGPMTQDMLFMQLRHHMTEAVNAGERAAGKPL----- 210
Db 170 ISGGDYRLALPVLRSDFEGDLTTAFNDM-----SRNLSIKDELLGEERAENQRLMSLMPE 225
Qy 211 -----PGARQVTVAFADLVGFTQLGEVVSABEL-----GHLAAGRLAGLARDL 252
Db 226 PVMQRYLDBEETIAQDHKNVTVIFADMMGLDELSEMLTSEELMVVVNDLTQFDAAAE 285
Qy 253 TAPPVWFIKTIGDAMLVCPDPAPLLDTVLKLV-----EYVD-----TNNFPRLRAGV 301
Db 286 ---GVDHVRTLHDGYLASCGLGVPRLDNVRVTNFAIEMDRIIDRHAESGHDLRLRAGI 342
Qy 302 ASGMVAVS-----RAGDWFSGPVNVASRV-TGVARPGAVLVADSVREALGDAPADG 353
Db 343 DTGSAAGSLVGRSTLAYDMMGSAVDVAVQVORGSPQPG-IYVTSRVHEVMQETILD----- 396
Qy 354 WSFAGPRRLRGIRGDVRLFRVR 375
Db 397 --FVAAGEVVGGRVETVWRLQ 416
```


Search completed: October 15, 2005, 03:24:12
Job time : 34.5955 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:14:10 ; Search time 94.9382 Seconds
(without alignments)
1742.576 Million cell updates/sec

Title: US-10-617-038-7
Perfect score: 1999
Sequence: 1 VTDHVREADDANIDLLGLD.....ARTSGGAAQDDLAGSSP 397

Scoring table: BLOSUM62

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA.*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep.*
 - 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
 - 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
 - 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1999	100.0	397	15 US-10-617-038-7	Sequence 7, Appli
2	308	15.4	405	14 US-10-156-761-10866	Sequence 10866, A
3	190.5	9.5	324	14 US-10-080-170-169	Sequence 169, App
4	190.5	9.5	324	16 US-10-080-170-169	Sequence 169, App
5	190.5	9.5	324	16 US-10-468-356-169	Sequence 169, App
6	185.5	9.3	173	16 US-10-425-115-342377	Sequence 342377,
7	168	8.4	185	14 US-10-032-189-90	Sequence 90, Appl
8	162	8.1	316	14 US-10-080-170-478	Sequence 478, App
9	162	8.1	316	16 US-10-080-170-478	Sequence 478, App
10	162	8.1	316	16 US-10-468-356-478	Sequence 478, App
11	159	8.0	356	15 US-10-369-493-11975	Sequence 11975, A

12	157	7.9	670	10	US-09-880-505-178	Sequence 178, App
13	157	7.9	670	13	US-10-051-643-178	Sequence 178, App
14	157	7.9	670	14	US-10-205-979-47	Sequence 47, Appl
15	157	7.9	722	13	US-09-880-505-174	Sequence 174, App
16	157	7.9	722	13	US-10-051-643-174	Sequence 174, App
17	151.5	7.6	566	20	US-11-097-143-12021	Sequence 12021, A
18	151	7.6	419	10	US-09-880-505-125	Sequence 125, App
19	151	7.6	419	13	US-10-051-643-125	Sequence 125, App
20	149	7.5	445	17	US-10-477-458-10	Sequence 10, Appl
21	148.5	7.4	194	14	US-10-032-189-91	Sequence 91, Appl
22	148.5	7.4	1690	20	US-11-097-143-19707	Sequence 19707, A
23	147.5	7.4	827	17	US-10-477-458-11	Sequence 11, Appl
24	147.5	7.4	1344	17	US-10-477-458-12	Sequence 12, Appl
25	144.5	7.2	619	14	US-10-354-358-116	Sequence 116, App
26	144.5	7.2	619	16	US-10-723-860-907	Sequence 907, App
27	142	7.1	447	17	US-10-477-458-9	Sequence 9, Appli
28	142	7.1	454	17	US-10-477-458-8	Sequence 8, Appli
29	142	7.1	908	18	US-10-840-512-199	Sequence 199, App
30	140.5	7.0	9234	10	US-09-942-025-13	Sequence 13, Appl
31	135	6.8	462	9	US-09-712-363-209	Sequence 209, App
32	133	6.7	1075	15	US-10-276-774-2153	Sequence 2153, Ap
33	132	6.6	690	20	US-11-097-143-7872	Sequence 7872, Ap
34	131.5	6.6	188	16	US-10-437-963-194003	Sequence 194003,
35	131.5	6.6	825	15	US-10-425-114-72756	Sequence 72756, A
36	131.5	6.6	895	17	US-10-762-107-6	Sequence 6, Appli
37	131.5	6.6	1300	15	US-10-282-122A-60085	Sequence 60085, A
38	131.5	6.6	2630	15	US-10-369-493-3693	Sequence 3693, Ap
39	131	6.6	8438	17	US-10-937-730A-4	Sequence 4, Appli
40	130	6.5	1076	20	US-11-097-143-5844	Sequence 5844, Ap
41	128.5	6.4	422	16	US-10-482-706-93	Sequence 93, Appl
42	128	6.4	1144	14	US-10-032-189-86	Sequence 86, Appl
43	128	6.4	1144	17	US-10-789-378-8	Sequence 8, Appli
44	127.5	6.4	736	15	US-10-425-114-70170	Sequence 70170, A
45	127.5	6.4	994	17	US-10-732-923-13512	Sequence 13512, A

ALIGNMENTS

RESULT 1
US-10-617-038-7
; Sequence 7, Application US/10617038
; Publication No. US20040057963A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Peter
; APPLICANT: Rosenkands, Ida
; APPLICANT: Stryhn, Anette
; TITLE OF INVENTION: Therapeutic TB Vaccine
; FILE REFERENCE: SS15AUSA
; CURRENT APPLICATION NUMBER: US/10/617,038
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: DK PA 2002 01098
; PRIOR FILING DATE: 2002-07-13
; PRIOR APPLICATION NUMBER: US 60/401,725
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis

Query Match	100.0%	Score	1999	DB	15	Length	397
Best Local Similarity	100.0%	Pred. No.	9.9e-172	Indels	0	Gaps	0
Matches	397	Conservative	0	Mismatches	0		
Qy	1	VTDHVREADDANIDLLGLDGLGTARAEAKLVLEQGITPDEIRATNPPLLLATRLHLV	60				
Db	1	VTDHVREADDANIDLLGLDGLGTARAEAKLVLEQGITPDEIRATNPPLLLATRLHLV	60				
Qy	61	GDGTGTVSAREISENYGVDELLELQVRAGVLARVDDPDVAVHMRADGFAAARQRFVEL	120				

```
Db 61 GDDGTYVSAREISENYGVDELLQVRQAVGLARVDDPDPAVVMRADGEGAAARQRFVEL 120
Qy 121 GLNPQVVLVVRVLAEGLSHAABANRYTALEIMRPGATELDTIAKGSQALVSQIVPLLGP 180
Db 121 GLNPQVVLVVRVLAEGLSHAABANRYTALEIMRPGATELDTIAKGSQALVSQIVPLLGP 180
Qy 181 MIQDLMFQOLRHMTEAVNAGERAAGKPLPGARQVTVAFADLVGFTOLGEVVSAAEELGH 240
Db 181 MIQDLMFQOLRHMTEAVNAGERAAGKPLPGARQVTVAFADLVGFTOLGEVVSAAEELGH 240
Qy 241 LAGRLAGLARDLTAPVNFIKTIGDVMLVCPDPAPLDITVLKLVVVVDITDNNFPLRAG 300
Db 241 LAGRLAGLARDLTAPVNFIKTIGDVMLVCPDPAPLDITVLKLVVVVDITDNNFPLRAG 300
Qy 301 VASGNVSRAGDWFSGSPVNVASRVTVGAPGAVLVADSVREALGDAPEADGFGQSFAGPR 360
Db 301 VASGNVSRAGDWFSGSPVNVASRVTVGAPGAVLVADSVREALGDAPEADGFGQSFAGPR 360
Qy 361 RLRGIRGVDVRLFRVRGATRTGSGGAQDDDLGSSP 397
Db 361 RLRGIRGVDVRLFRVRGATRTGSGGAQDDDLGSSP 397

RESULT 2
US-10-156-761-10866
; Sequence 10866, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10866
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10866

Query Match 15.4%; Score 308; DB 14; Length 405;
Best Local Similarity 31.2%; Pred. No. 5.9e-19;
Matches 100; Conservative 44; Mismatches 142; Indels 34; Gaps 9;

Qy 42 PDEIRATNPPLLATRLHVLGDDGTYVSAREISENYGVDELLQVRQAVGLARVDDPDPAV 101
Db 54 PEE----NPLALRLQLILGAERRY-TFPQARSAGVSMELASRFRWANGFADICQAKAL 108

Qy 102 VHMRADEGAAAARQRFVELG-LNPQVVLVVRVLAEGLSHAABANRYTALEIM---RPG 157
Db 109 T--EADVLALRLAGLVEAGLLSEAMVQVARSQTQTARLAEWQIDGFLGLETPPEPFG 166

Qy 158 ATELDIAKGSQALVSQIVPLLGPMTQDMLFMQLRHMTEAVNAGERAAGKPLPCA---- 213
Db 167 MTRTEV-----TYPLVELLLPELEEFLLVYVWRQLAA-----ATGRVVOAGDDEE 211

Qy 214 ---RQTVAFADLVGFTOLGEVVSAAEELGHLAGRLAGLARDLTAPPVW-FIKTIGDVML 269
Db 212 MYDRRLAVGFADLVGFTRLTRMEEEELGELVEAFETTAADLVAAHGGLIKTLGDEVLY 271

Qy 270 VCPDPAPLDITVLKLVVVVDITDNNFPLRAGVAGSNVSRAGDWFSGSPVNVASRVTVGAR 329
Db 272 AADDAGVAAEIALRIETMANDETPELRVGFAGFTVTRMGDVFGTTVNLASRLTSIAP 331
```

```
Qy 330 PGAVLVADSVREALGDAPEA 349
Db 332 RDAVLVDGAFABELTRTREAA 351

RESULT 3
US-10-080-170-169
; Sequence 169, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495-0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 169
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
US-10-080-170-169

Query Match 9.5%; Score 190.5; DB 14; Length 324;
Best Local Similarity 25.9%; Pred. No. 1.8e-08;
Matches 95; Conservative 46; Mismatches 139; Indels 87; Gaps 15;

Qy 42 PDEIRATNPPLLATRLHVLGDDGTYVSAREISENYGVDELLQVRQAVGLARVDDPDPAV 101
Db 28 PDEDTVTVPAPALIRRSVS-----MPNAAQWLHTTNRSPRLVAMVR----- 69

Qy 102 VHMRADEGAAAARQRFVELGPNQVVLVVRVLAEGLSHAABANRYTALEIMRPGATEL 161
Db 70 -----RARRL--LPGDDP-----FGDPLS-----TAGEGGPRAAARAA 100

Qy 162 DIAKGSQALVSQIVPLLGPMTQDMLFMQLRHMTEAVNAGERAAGKPLPGARQVTVAPA 221
Db 101 DRLLGDRGAASREVSLSVLQWQAL-----TEAI-----ARRPV--NPEVTLVFT 143

Qy 222 DLVGFTOLGEVVSAAEELGHLAGRLAGLARDLTAPPV-----WFIKTIGDVMLVCPDPAI. 277
Db 144 DLVGFS--GWSLQAGDEATLA--LLRQVARAVESPLLDAGGHIVKRMGGIWAFFRDPVA 200

Qy 278 LDTVLKLVFV---VDTDNNFRLRAGVASGNVSRAGDWFSGSPVNVASRVTVGARPGAVL 334
Db 201 VOAVLAATEAMKSVEVGCTPRIRVGIHTGRPQLAADMLGVDVNIARVMERATKGGIM 260

Qy 335 VADSVREALGDAPEADGFGQSFAGPRRLR-----GIRGDVRLFRVRGATRTGSGG 385
Db 261 ISGPTLDLI---PQSLDKELGELT-TRRVKPKMFTSKFTGIPDPMVIYRIKARRELTA 316

Qy 386 AAQDDDL 392
Db 317 TAQTNLS 323

RESULT 4
US-10-080-170-169
; Sequence 169, Application US/10080170
; Publication No. US20040121322A9
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495-0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
```


Qy	211	-----PGARQVTVAFADLVTGFTOLGEVWSAEEL-----GHUAGRLAGLAGIARDL	252
Db	476	PVMORYLDGEETIAQDHKNVTVIFADMGLDELDSRMLTSEELMVVVNDLTQFDAAAEEL	535
Qy	253	TAPPVWFKTIGDAVMLYCPAPALDITVLKLV-----EVVD-----TDNNFPRLRACGV	301
Db	536	---GVDHVRTLDHGYSACGLGVPLRDNVRRVTNFAIEMDRDIIRHAESGHDLRLRAGI	592
Qy	302	ASGMAYS-----RAGDWFGSPVNVASRV-TGVARPGAVLVADSRE-----ALG	344
Db	593	DTGSNAASGLVGSRSLTAYDMWGSADVAVANQVRGSGPQG-IYTVSRHVHEVMQETLDFVAAG	651
Qy	345	DAPEADGFQ--WSFAGPRR	361
Db	652	EWGERGGVETVWRLOGHRR	670

RESULT 15

```

US-09-880-505-174
; Sequence 174, Application US/09880505
; Publication No. US20030007976A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L.J.
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
; FILE REFERENCE: 11000.1007C2
; CURRENT APPLICATION NUMBER: US/09/880,505
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 09/324,542
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: US 08/997,080
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 174
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
; US-09-880-505-174

```

Query Match	7.9%;	Score 157;	DB 10;	Length 722;
Best Local Similarity	23.9%;	Pred. No. 5.8e-05;		
Matches	105;	Conservative 55;	Mismatches 159;	Indels 120; Gaps 20;
Qy	14	DLGLDGLGGTA-----	RAERAKLVEMLEEQGITPDEI-----	RATNPP 51
Db	313	DTGMDGTGETILVGPDNLMRSDSKLFRENREKFLADVVEGG-TPEVADESVDRRGTTLV 371		
Qy	52	LLLATRHLV-----	VYSAREISENTGVDELLIQRVQRAVGAVLRVDDPDVV 102	
Db	372	QPVTRSVEEAQRNGTGTITIEDDYLGHQALQAYSPVDLPGLHWI-----	VAKID-----	421
Qy	103	HMRADGEAARAQRVELGLGNPDQVLRVVRVLAEGLSHAABEMRYTALTAIMRPGATELD 162		
Db	422	---TDEAFAPVAQPTR-----	TLVLSVIIIFGVSLAAMLLARLFVRPIRRLQAGAAQ 471	
Qy	163	IAKSQALVSQIVP-----	LGPMIQDLMFQLRHHMETEAVNAGERAAQKPL-----	210
Db	472	ISGGDYRLALPVLRFDEFGDITTAFTNDM-----	SNLSIKDELLGEERAENQRLMSLMPE 527	
Qy	211	-----	PGARQVTVAAADILVGFQTLGEVVSABEL-----	GHLAGRLAGLARDL 252
Db	528	PVMQRYLDGESTIAQDHKNVTVIADMMGLDELSRMLTSBELMVVNDLTQFPDAAEEL 587		
Qy	253	TAPPWFITKTCDAVMLVCPDPAPLLDTVLKLV-----	EVVD-----	TDNMFPLRAGV 301
Db	588	---GVDHVRTLHDGVLACGGLGVPRLDNVRTVTFVFAEMDRIIIDRHAAESCHDLRLAGI 644		
Qy	302	ASGMAYS-----	RAGDWFCSPPVNVASRV-TGVARPGAVILVADSVE-----	ALG 344

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:59:30 ; Search time 22.1523 Seconds
(without alignments)
1724.342 Million cell updates/sec

Title: US-10-617-038-7

Perfect score: 1995
Sequence: 1 VTDHVREADDANIDLLGL.....ATRTGGGAQDDDLGSSP 397

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1996	99.8	397	H70753	hypothetical prote
2	320	16.0	387	S42625	adenylate cyclase
3	271	13.6	403	OYRKO	adenylate cyclase
4	245	12.3	378	E70786	hypothetical prote
5	190.5	9.5	324	A87084	conserved hypothet
6	162	8.1	316	B70620	hypothetical prote
7	157	7.9	472	AD1889	adenylate cyclase
8	156.5	7.8	520	H95884	probable regulator
9	152	7.6	1155	AC2426	adenylate cyclase
10	151.5	7.6	418	G70558	hypothetical prote
11	151.5	7.6	1159	B95370	probable adenylate
12	150.5	7.5	352	T10905	adenylate cyclase
13	149.5	7.5	1118	C95385	probable adenylate
14	145	7.3	1110	I59370	guanylate cyclase
15	144.5	7.2	619	S23097	guanylate cyclase
16	142	7.1	730	A70680	hypothetical prote
17	142	7.1	1313	T30548	adenylate cyclase
18	140.5	7.0	546	AE1899	adenylate cyclase
19	140.5	7.0	555	D7536	methyl-accepting c
20	138	6.9	1331	T18310	receptor-adenylate
21	137	6.9	1235	D32433	VSG expression sit
22	136.5	6.8	735	AC1946	adenylate cyclase
23	136	6.8	1057	OYRTR	atrial natriuretic
24	136	6.8	1057	I55319	guanylyl cyclase A
25	136	6.8	1057	I57963	natriuretic peptid
26	135.5	6.8	540	T34187	hypothetical prote
27	135	6.8	462	A70518	probable lipJ prot
28	135	6.8	1380	T18309	receptor-adenylate
29	134.5	6.7	619	OYB070	guanylate cyclase

30	134.5	6.7	619	1	OYRTB1	guanylate cyclase
31	134	6.7	226	2	C95421	probable adenylate
32	133	6.7	1061	1	OYHUAR	natriuretic peptid
33	133	6.7	1277	2	T30876	adenylate cyclase
34	132.5	6.6	637	2	B95878	probable adenylate
35	132.5	6.6	1108	2	B55915	guanylate cyclase
36	132.5	6.6	1202	2	T17197	adenylate cyclase
37	132	6.6	1103	2	JC5581	guanylate cyclase
38	131	6.6	1057	1	OYMSAR	atrial natriuretic
39	130.5	6.5	1056	2	T33167	hypothetical prote
40	130	6.5	1108	2	I59385	guanylate cyclase
41	130	6.5	1122	2	T28130	hypothetical prote
42	129.5	6.5	1321	2	T30537	adenylate cyclase
43	129	6.5	1068	2	T42382	guanylate cyclase
44	128.5	6.4	422	2	H70914	hypothetical prote
45	128.5	6.4	505	2	T18933	hypothetical prote

ALIGNMENTS

RESULT 1

H70753

hypothetical protein Rv1264 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: H70753

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;

Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Bartell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: H70753

A:Status: Preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-397 <COL>

A:Cross-references: UNIPROT:Q11055; GB:Z77137; GB:AL123456; NID:g3261593; PIDN:CAB00890.1

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: Rv1264

Query Match 99.8%; Score 1996; DB 2; Length 397;

Best Local Similarity 99.7%; Pred. No. 2.le-125;

Matches 396; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VTDHVREADDANIDLLGLGCTARAEAKLVENLLEQGITDEIRATNPPLLATRHVLV 60

DB 1 MTDHVREADDANIDLLGLGCTARAEAKLVENLLEQGITDEIRATNPPLLATRHVLV 60

QY 61 GDDGTYVSAREISENKGVDLELLQVRQAVGLARVDDPDVAVHMRADGEAAARQRFVEL 120

DB 61 GDDGTYVSAREISENKGVDLELLQVRQAVGLARVDDPDVAVHMRADGEAAARQRFVEL 120

QY 121 GLNPQVWLVRVLAEGLSHAAEAMRYTALEAIMRPGATELDAKGSQALVSQIVPLLGP 180

DB 121 GLNPQVWLVRVLAEGLSHAAEAMRYTALEAIMRPGATELDAKGSQALVSQIVPLLGP 180

QY 181 MTQDMLFMQLRHMMETEAVNAGERAGKPLPGARQVTVAFADLVGFTQLGVEVVAEELGH 240

DB 181 MTQDMLFMQLRHMMETEAVNAGERAGKPLPGARQVTVAFADLVGFTQLGVEVVAEELGH 240

QY 241 LAGRLAGLARDLTAPPVWFKITGDVAVMLVCPDPAFLDITVLKLVVEVDTNNFPLRAG 300

DB 241 LAGRLAGLARDLTAPPVWFKITGDVAVMLVCPDPAFLDITVLKLVVEVDTNNFPLRAG 300

QY 301 VASGMVSRAGDMFGSPVNVASRVTVGARPGAVLVADSVREALGDAPEADGQWFSAGPR 360

DB 301 VASGMVSRAGDMFGSPVNVASRVTVGARPGAVLVADSVREALGDAPEADGQWFSAGPR 360

QY 361 RLRGIRGVDVLRFRVRGATRTGSGGAQDDDLGSSP 397

|||||

Db 361 RLIRGIRGVDLPRVRRGATRTGSGGAQDDDLGSSP 397

RESULT 2

adenylate cyclase (EC 4.6.1.1) - Streptomyces coelicolor (fragment)
C:Species: Streptomyces coelicolor
C:Date: 07-Sep-1994 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
C:Accession: S42625; 742035
R:Danchin, A.; Pidoux, J.; Krin, E.; Thompson, C.J.; Ullmann, A.
FEMS Microbiol. Lett. 114, 145-152, 1993
A:Title: The adenylate cyclase catalytic domain of Streptomyces coelicolor is carboxy-terminally truncated.
A:Reference number: S42625; MUID:94109670; PMID:8282183
A:Accession: S42625
A:Molecule type: DNA
A:Residues: 1-387 <DAN>
A:Cross-references: EMBL:X74768; NID:G433913; PIDN:CAA52780.1; PID:G456210
C:Genetics:
A:Gene: cya
C:Superfamily: Brevibacterium liquefaciens adenylate cyclase
C:Keywords: phosphorus-oxygen lyase

Query Match 16.0%; Score 320; DB 2; Length 387;
Best Local Similarity 29.5%; Pred. No. 5.1e-14;
Matches 121; Conservative 47; Mismatches 192; Indels 50; Gaps 13;

QY 5 VREADDANIDLLGDLGGTARAERAKLVLEWLLSQGITPDEIRATNPPLLLATRHVLV-GDD 63
Db 1 IGKAAAGVTVDITGSGAGDGRVDPEP-----APDSADPGSDPLALRLLEGILGAE 50

QY 64 GTYVSAREISENYGVDELLELQVORAVGLARVDDPDVAVHMRADGEAAARQRFVELG-L 122
Db 51 RRY-TFPOAASAGVSMELASFWAMGFADIGQAKAIT--EADVLAURLAGLVEAGLL 107

QY 123 NPDQVVLVVRVLAEGSLHAAEAMRYTALAEINM---RPGATELDIAKGSQALVSIQVPLLG 179
Db 108 SEAMAVQVARSTGQTARLAEQIDSFLEGLTEPPEPGWTRTEV-----TYPIVEELL 160

QY 180 PMIQDMLFMQLRHMMETAVNAGE--RAAGKPLFGARQVTVAFADLVGFTOLGEVVSAAE 237
Db 161 PELQSFVLYVWRQL---AASAGRVIQAGDDEEMVDRRLVAGFADLVGFTLRLRMEEEE 217

QY 238 LCHLAGRLAGLARDLTAPPVW-FIKTIGDAVNLVCPDPAPLLDTVLKLVVVVDITDNNPFR 296
Db 218 LGELVEAETTSADLVARGGLVKTGLDEVLYAADDAGTAAEIALLLVETMAHDETNP 277

QY 297 LRAGVASGMVSRAGDWFSGSPNVASRVTVGVARPGAVLVADSVREAL---GDAPADGQF 353
Db 278 LRVGIAFGVTTRMGDFVCTTNLASRLTSTAPKDAVLVDVTAFAEELIRTRDAPASEAAA 337

QY 354 WSFAGPRRLRGIRGVDLPRVR-----RG-----ATRTGSGGAA 387
Db 338 ABEAAAEKEGEPVYRFALQPMQWRPVRLGVVPEWMLLRRDGGGGEA 387

RESULT 3

OYFKQ

adenylate cyclase (EC 4.6.1.1) - Brevibacterium liquefaciens
C:Species: Brevibacterium liquefaciens
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: S15273
R:Peters, E.P.; Wilderapin, A.F.; Wood, S.P.; Zvelebil, M.J.J.M.; Sezer, O.; Danchin, A.
Mol. Microbiol. 5, 1175-1181, 1991
A:Title: A pyruvate-stimulated adenylate cyclase has a sequence related to the fes/fps of Bacillus subtilis.
A:Reference number: S15273; MUID:92065813; PMID:1683468
A:Accession: S15273
A:Molecule type: DNA
A:Residues: 1-403 <PET>
A:Cross-references: UNIPROT:P27580; EMBL:X57541; NID:G48831; PIDN:CAA40760.1; PID:G48832
C:Genetics:
A:Gene: cya
C:Superfamily: Brevibacterium liquefaciens adenylate cyclase
C:Keywords: CAMP biosynthesis; dimer; phosphorus-oxygen lyase

Query Match 13.6%; Score 271; DB 1; Length 403;
Best Local Similarity 26.3%; Pred. No. 9.6e-11;
Matches 106; Conservative 64; Mismatches 181; Indels 52; Gaps 15;

QY 4 HVREADDANIDLLGDLGGTARAERAKLVLEWLLSQGITPDEIRATNPPLLLATRHVLVGD 62
Db 27 HAPATPAESDPIL-----ELAEAMEGFLRAIPAHPTPEAVRDTVASL---EKRLIGG 74

QY 63 DGTYSAREISENYGVDELLELQVORAVGLARVDDPDVAVHMRADGEAAARQRFVELGL 122
Db 75 QREF-RRREVASEAGVSLHARKLWRAIGFPFLSD-DEVFFTEADKKALGTVMGVMVRGA 132

QY 123 NPDQVVLVVRVLAEGSLHAAEAMRYTALAEINM---RPGATELDIA-KGSOALVSIQVPLIGPM 181
Db 133 LTEETAI---SLMRVGOMTDMVVWQIEALVEDMIANQNLSQDRAQLQLSLPEIIPA 189

QY 182 IQDMLFMQLRHMMETEA-----VNAGERAAGKPLFGARQVTVAFADLVGF 226
Db 190 IEDLLLYSWRRQLNSAVHMRALRVETGVAAVQNDRGEDDGGTPLPLAR--AVGFADLVSY 247

QY 227 TOLGEVVSAREISENYGVDELLELQVORAVGLARVDDPDVAVHMRADGEAAARQRFVELGL 285
Db 248 TSLSRMMNERTLAQLVORFEAKCAEIIISVGGRLVKTIGDEVLYVAETPQAGATLSLS 307

QY 286 EVVTDNNFPLRAGVASGMVSRAGDWFSGSPNVASRVTVGVARPGAVL---VADSVRE 341
Db 308 RELAKDELFPQTRGAVVWGRLLSLGDIYGTVMNAARLTSLAEFGTVLTDAITANTUR- 366

QY 342 ALGDAPEADGQWGFAGPRRLRGIRGVDLPRVRRGATRTGSG 384
Db 367 ---NDA---RFVLTAQETAVRGF-GDIQPVLSAG---EGAG 399

RESULT 4

E70786

hypothetical protein Rv2212 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: E70786
R:Coile, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: E70786
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-378 <COL>
A:Cross-references: UNIPROT:Q10400; GB:Z70283; GB:AL123456; NID:G3261561; PIDN:CAA94255.1
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv2212

Query Match 12.3%; Score 245; DB 2; Length 378;
Best Local Similarity 27.1%; Pred. No. 4.8e-09;
Matches 109; Conservative 61; Mismatches 152; Indels 80; Gaps 18;

QY 9 DDANIDLLGDLGGTAR-AERAKLVLEWLLSQGITPDE-IRA-----TNP 50
Db 3 DLSDFDAL--EAAGIANPREAGLLTYLDELGFTVEENVQARRGLRGLAGDVLLWSGP 60

QY 51 PL-LLATRHVLVGDGTYSAREISENYGVDELLELQVORAVGLARVDDPDVAVHMRADGE 109
Db 61 PIYTLAT-----AADELGLSADVARAWSLGLT-VAGPDVPTLSQADV 104

QY 110 AAARQRFVELGLNPDQVVLVVRVLAEGSLHAAEAMRYTALAEINMPCA-----TE 160
Db 105 ALATWALKAL-VGEDGAFGLRVLGTMARLAEAE-----STMIRAGSPNIQMTHTHDE 158

QY 161 LDIKGSQALVSIQVPLIGPMIQDMLFMQLRHMMET-----EAVNAGBRAAGKPLPGARQ 215

Db 159 LATAAYRA-AAEFVRPRIGALIDTV-----HRHLSAARYFEV-IGD TSA-----S 204

Qy 216 VT--VAFADLVGFTQLGVEVVSABELGHLAAGLAGLARDLT-APPVVMFIKTIGDVMVLVCP 272

Db 205 VTCIGIGFADLSSFTALTQALTPAQLQDLLETFDAVTVVHADGGRLVKFIGDVMVWSS 264

Qy 273 DPAPLDTVLKVEVVDVTDNNFRLRAGVAGSMVSRAGDWFSGSPVNVASRTVGVARPGA 332

Db 265 SPERLVRAAVDLVDHPGARAELQVRAGLAYGTVLALNGDYFGNPNVLAARLVAAAAPGQ 324

Qy 333 VLVADSVREALGDAPEADGFQWFSFAGPRRLRGIRGVDVRLFRV 374

Db 325 ILAAQLRMLPDWPALE-----AHGLTLKGFADPVMAPFL 360

RESULT 5

A87084

conserved hypothetical protein ML1399 [imported] - Mycobacterium leprae

C:Species: Mycobacterium leprae

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C:Accession: A87084

R: Cole, S.T.; Eiglmeyer, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho

R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,

eam, M.A.; Rutherford, K.M.

Nature 409, 1007-1011, 2001

A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S

A:Title: Massive gene decay in the leprosy bacillus.

A:Reference number: A86909; MUID:21128732; PMID:11234002

A:Accession: A87084

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-324 <STO>

A:Cross-references: UNIPROT:Q9CC18; GB:AL450380; NID:gl3093284; PIDN:CAC31780.1; GSPDB:G

C:Genetics:

A:Gene: ML1399

Query Match 9.5%; Score 190.5; DB 2; Length 324;

Best Local Similarity 25.9%; Pred. No. 1.7e-05;

Matches 95; Conservative 46; Mismatches 139; Indels 87; Gaps 15;

Qy 42 PDEIRATNPILLATRHLYGDDGTGYVSAREISENYGVDELQVRQVAVGLARVDDPDVAV 101

Db 28 PDEITVTVFAPALIRSSVS-----WPNAAQWLHTNRSPLRVAMVR----- 69

Qy 102 VHMRADEAAARQREVELGNPDQVVLVVRVLAELGSHAAEAMRYTALEAMRPCATBL 161

Db 70 -----RARLL--LPGDPD-----FGDPLS-----TAGEGPRAAARAA 100

Qy 162 DIAKGSQALVSQIVPLLGPMIQDMLPMQLRHMETEAENVAGERAAGKPLPGARQVTVAPA 221

Db 101 DRLLGDRGAASREVLSVLQVMQAL-----TEAI-----ARRPV--NPEVTLVFT 143

Qy 222 DLVGFQTLGVEVVSABELGHLAAGLAGLARDLTAPPV-----WFIKTIGDVMVLVCPDPAPL 277

Db 144 DLVGFQ--GWSLQAGDEATLA--LLRQVARAVASPLLDAGGHVYKRGDGDIMAVFRDPVSA 200

Qy 278 LDTVLKLVFV---VDTDNNFRLRAGVAGSMVSRAGDWFSGSPVNVASRTVGVARPGVL 334

Db 201 VQAVLAATAMKSVEGGYTPRIRVGIHTGRQORLAADWLGVVDVNIARVMERATKGGIM 260

Qy 335 VADSVREALGDAPEADGFQWFSFAGPRRLR-----GIRGDVRLFRVRGATRTGSGG 385

Db 261 ISGPTLDLI---PQSDKLKELGII--TRVRKPMFTSKFTGIPDMVVIYRIKARRELTASDE 316

Qy 386 AAQDDDL 392

Db 317 TAQTNLSL 323

RESULT 6

B70620

hypothetical protein Rv1647 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: B70620

R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

R.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Genes, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: B70620

A:Status: Preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-316 <COL>

A:Cross-references: UNIPROT:P94982; GB:Z85982; GB:AL123456; NID:g3261718; PIDN:CAB0664.1

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: Rv1647

Query Match 8.1%; Score 162; DB 2; Length 316;

Best Local Similarity 26.5%; Pred. No. 0.0013;

Matches 62; Conservative 27; Mismatches 63; Indels 82; Gaps 9;

Qy 196 TEAVNAGERAACKPLPGARQVTVAFADLVGF-----TOLGEVVSAREEL----- 238

Db 117 TEAVSR-----RPAPEVTLVFTDLVGFSTWSLHAGDDATLTLRQVARAVESPLLD 168

Qy 239 --GHLAAGLAGLARDLTAPPVVMFIKTIGDVMVLVCPDPAPLDTVL---KLVEVVDVTDNN 293

Db 169 AGGHI-----VKRLGDGINAVFRNPTVALRAVLVAQDAVKSLEVOGY 210

Qy 294 FPLRAGVAGSMVSRAGDWFSGSPVNVASRTVGVARPGAV-----LVADSVREALGD 345

Db 211 TPRMIGITGRPQRLAADWLGVNIAARVMERATKGGIMISQPTLDLIPQSELDALG- 269

Qy 346 APEADGFQWFSFAGPRRLR-----GIRGDVRLFRVRGATRTGSGAAQDD 390

Db 270 -----VVARRVRKPVFASKPTGIPDLAIYRI-----KTVSESTAADN 307

RESULT 7

AU1889

adenylate cyclase [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C:Accession: AD1889

R: Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,

Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anab

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AD1889

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-472 <KUR>

A:Cross-references: UNIPROT:Q8YZ26; GB:BA000019; PIDN:BA872619.1; PID:gl7130007; GSPDB:G

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all0661

Query Match 7.9%; Score 157; DB 2; Length 472;

Best Local Similarity 25.2%; Pred. No. 0.0044;

Matches 98; Conservative 59; Mismatches 150; Indels 82; Gaps 21;

Qy 17 LCDLGGTAERAKLVWLEOGITPDEIRATNPPELLLATRHLYGDDGTGYVSAREISENY 76

Db 122 LGKIQPSGKGFREAAAGRFLEGGV-PNN---TNP--LLIRRVNVSEKPT-----LAQLL 169

Qy 77 GVDLELLQVRQVAVGLARVDD-----PDVAVHMRADGEAAARQRFVELGLNPDQVVLV 131

Db 170 AERVLMQ-----GLEIGDWFTKETAGVNRKQVEYTLAAILFVLFL---EGILIF 220

[illegible]

A:Map position: 4q31.3-4q33
C:Superfamily: soluble guanylate cyclase; guanylate cyclase catalytic domain homology
C:Keywords: heterodimer; phosphorus-oxygen lyase
F:373-607/Domain: guanylate cyclase catalytic domain homology <GCC>

Query Match		7.2%	Score 144.5;	DB 1;	Length 619;
Best Local Similarity		22.7%	Pred. No. 0.041;		
Matches		69;	Conservative 46;	Mismatches 106;	Indels 83; Gaps 14;
QY	119	ELGLNPQVVLVVRVLAEGLSHAAEAMRYTALAEIMRPGATELDIAKGSQALVSQIVPLL	178		
Db	368	ELEILTRLQLTLR-----ALE-----DEKKKTDLLYSVLP--	399		
QY	179	GPMIQDMLFMQLRHMMETEA VNAGERAAAGKPLPGAR--QVTAFADLVGFTOLGEVVSAE	236		
Db	400	-PSVAN-----ELRHK-----RPVPAKRYDNVTILFSGIVGFNAF---CSKH	437		
QY	237	ELGHLAAGRLAGLARDLTA-----PPVWFIKTIGDAVMLVCPDPAP-----LL	278		
Db	438	ASGEGAMKIVNLLNDLYTRFDLTLSRKKNPFVYKVTGVGDYKNTVSGLPEPCIHARSIC	497		
QY	279	DTVLKLVEV---VDTDNFPRLRAGVASGMVVSRA-GD-----WFGSPVNVASRVTVGA	328		
Db	498	HLALDMWEIAGQVOVDGESVQITIGTGEVTVGQRMPCRYCLEGNTVNLTSRTETTG	557		
QY	329	RPGAVLVADSVREALGDAPEAD-GFQWSFAGPRRLRGIRGVDVRLFRVRRGATRTGSGAA	387		
Db	558	EKGKINVSEYTYRCLMSPENSDDPFHLEHRGPVSMKGKKEPMQVWFLSR--KNTGTETETK	615		
QY	388	ODDD	391		
Db	616	ODDD	619		

Search completed: October 15, 2005, 04:38:24
Job time : 23.1523 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:24:25 ; Search time 88.9255 Seconds
(without alignments)
2286.133 Million cell updates/sec

Title: US-10-617-038-7

Perfect score: 1999

Sequence: 1 VTDHVREADDANIDLLGLD.....ATRTGGGAQDDDLGSSP 397

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1996	99.8	397	1	YC64_MYCTU	Q11055 mycobacteri
2	1998	99.4	397	2	Q7U096	Q7U096 mycobacteri
3	1472.5	73.7	385	2	Q73X03	Q73X03 mycobacteri
4	326	16.3	399	1	CYAA_STRGR	Q9wxc3 streptomyc
5	317	15.9	381	1	CYAA_STRCO	P40135 streptomyc
6	308	15.4	405	2	Q82I29	Q82I29 streptomyc
7	271	13.6	403	1	CYAA_BRELI	P27580 brevibacter
8	245	12.3	378	1	YM12_MYCTU	P64265 mycobacteri
9	245	12.3	378	1	YM35_MYCBO	P64266 mycobacteri
10	190.5	9.5	324	2	Q9CC18	Q9CC18 mycobacteri
11	186.5	9.3	481	2	Q899E3	Q899E3 rhizobium 1
12	172	8.6	412	2	Q7WV9	Q7WV9 alcaligenes
13	166.5	8.3	322	2	Q740J3	Q740J3 mycobacteri
14	163	8.2	527	2	Q89SW2	Q89SW2 bradyrhizob
15	162	8.1	316	2	P94982	P94982 mycobacteri
16	162	8.1	316	2	Q7TZU4	Q7TZU4 mycobacteri
17	162	8.1	327	2	Q8VJ25	Q8VJ25 mycobacteri
18	159	8.0	356	2	Q98GN7	Q98GN7 rhizobium 1
19	157	7.9	472	2	Q8YZ26	Q8YZ26 anabaena sp
20	156.5	7.8	520	2	Q92WJ5	Q92WJ5 rhizobium m
21	155.5	7.8	654	2	Q8DGN6	Q8DGN6 synecococc
22	155	7.8	640	2	Q6W199	Q6W199 rhizobium s
23	153	7.7	424	1	CYAA_STIAU	P40137 stigmatella
24	153	7.7	1105	2	Q6N2P9	Q6N2P9 rhodospheudo
25	153	7.7	2155	2	Q9HEP4	Q9HEP4 blumeria gr
26	152	7.6	1155	2	P94183	P94183 anabaena sp
27	152	7.6	1155	2	Q8YMH0	Q8YMH0 anabaena sp
28	151.5	7.6	443	1	CYAA_MYCTU	Q30820 mycobacteri
29	151.5	7.6	513	2	Q9V8D9	Q9V8D9 drosophila
30	151.5	7.6	1159	2	Q92YLO	Q92YLO rhizobium m
31	150.5	7.5	352	1	CYAB_STIAU	P40138 stigmatella

32	149.5	7.5	634	2	Q92M76	Q92M76 rhizobium m
33	149.5	7.5	1112	2	Q89FJ3	Q89FJ3 bradyrhizob
34	149.5	7.5	1118	2	Q68M96	Q68M96 rhizobium m
35	149.5	7.5	1118	2	Q92Y97	Q92Y97 rhizobium m
36	149	7.5	471	2	Q98GL5	Q98GL5 rhizobium 1
37	148.5	7.4	510	2	Q89FF4	Q89FF4 bradyrhizob
38	148.5	7.4	1690	2	Q9VXU0	Q9VXU0 drosophila
39	148.5	7.4	1708	2	O16039	O16039 drosophila
40	147	7.4	658	2	Q89ET6	Q89ET6 bradyrhizob
41	147	7.4	1096	2	O97350	O97350 trypanosoma
42	146.5	7.3	581	2	Q98LH2	Q98LH2 rhizobium 1
43	145.5	7.3	546	2	O7Q9X4	O7Q9X4 anopheles g
44	145	7.3	1110	1	CYGX_RAT	P51839 rattus norv
45	144.5	7.2	445	2	Q89BY7	Q89BY7 bradyrhizob

ALIGNMENTS

RESULT 1

YC64_MYCTU
ID YC64_MYCTU STANDARD; PRT; 397 AA.
AC Q11055;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Hypothetical protein RV1264/MT1302.
GN OrderedLocusNames=RV1264, MT1302; ORFNames=MTCY50.18C;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holtroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrall B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
Fraser C.M.;
RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains".
RT Laboratory strains".
RL J. Bacteriol. 184:5479-5490(2002).
CC -!- SIMILARITY: Belongs to the adenyl cyclase class-4/guanylyl
cyclase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; BX842576; CAB00890.1; -.

```

DR EMBL; AE000516; AAK45561.1; -.
DR PIR; H70753; H70753.
DR HSP; Q02846; 1ANL.
DR TIGR; MT1302; -.
DR Tuberculin; RV1264; -.
DR InterPro; IPR001054; G_cyclase.
DR Pfam; PF00211; Guanylate_cyc; 1.
DR SMART; SM00044; CYCC; 1.
DR PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR Complete proteome; Hypothetical protein.
KW DOMAIN 217 325 Guanylate cyclase.
SQ SEQUENCE 397 AA; 42232 MW; F6C212A181DBSAD3 CRC64;

Query Match 99.8%; Score 1996; DB 1; Length 397;
Best Local Similarity 99.7%; Pred. No. 2e-118;
Matches 396; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VTDHVREADDANIDLLGLGCTARAEAKLVEMLEOGITPDEIRATNPPLLATRLHV 60
DB 1 MTDHVREADDANIDLLGLGCTARAEAKLVEMLEOGITPDEIRATNPPLLATRLHV 60

QY 61 GDDGTYSAREISENYGVDELLELQVQRAVGLARVDDPDVAVHMRADGEEAARQRFVEL 120
DB 61 GDDGTYSAREISENYGVDELLELQVQRAVGLARVDDPDVAVHMRADGEEAARQRFVEL 120

QY 121 GLNPDQVVLVVRVLAEGLSHAAEAMRYTALBAIMRPGATELIDIAKGSQALVSQIVPLLGP 180
DB 121 GLNPDQVVLVVRVLAEGLSHAAEAMRYTALBAIMRPGATELIDIAKGSQALVSQIVPLLGP 180

QY 181 MIQDMLFMQLRHMMETEAUNAGERAAGKPLPGARQVTVAFADLVGFTQLGWVSAEELGH 240
DB 181 MIQDMLFMQLRHMMETEAUNAGERAAGKPLPGARQVTVAFADLVGFTQLGWVSAEELGH 240

QY 241 LAGRLAGLARDLTAPPVWFITIGDAVNLVCPDPAPLLDTVLKLVVVDTDNFPRLRAG 300
DB 241 LAGRLAGLARDLTAPPVWFITIGDAVNLVCPDPAPLLDTVLKLVVVDTDNFPRLRAG 300

QY 301 VASGMVSRAGDWFPGSPVNVASRVTVGVARPGAVLVADSVREALGDAPADGFGQSFAGPR 360
DB 301 VASGMVSRAGDWFPGSPVNVASRVTVGVARPGAVLVADSVREALGDAPADGFGQSFAGPR 360

QY 361 RLIRGIRGVRLFRVRRGATRTGSGGAAQDDDLGSSP 397
DB 361 RLIRGIRGVRLFRVRRGATRTGSGGAAQDDDLGSSP 397

RESULT 2
QY096 PRELIMINARY; PRT; 397 AA.
AC QY096;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE ADENYL CYCLASE (ATP PHOSPHATASE-LYASE) (ADENYLATE CYCLASE) (EC 4.6.1.1).
GN OrderedLocusNames=Mb1295;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=2709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eigmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248338; CAD94156.1; -.
DR GO; GO:0004383; F:guanylate cyclase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.

```

```

DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001054; G_cyclase.
DR Pfam; PF00211; Guanylate_cyc; 1.
DR SMART; SM00044; CYCC; 1.
DR PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR Complete proteome; Lyase.
SQ SEQUENCE 397 AA; 42179 MW; E3D8AD1E31C1F563 CRC64;

Query Match 99.4%; Score 1988; DB 2; Length 397;
Best Local Similarity 99.5%; Pred. No. 6.3e-118;
Matches 395; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VTDHVREADDANIDLLGLGCTARAEAKLVEMLEOGITPDEIRATNPPLLATRLHV 60
DB 1 MTDHVREADDANIDLLGLGCTARAEAKLVEMLEOGITPDEIRATNPPLLATRLHV 60

QY 61 GDDGTYSAREISENYGVDELLELQVQRAVGLARVDDPDVAVHMRADGEEAARQRFVEL 120
DB 61 GDDGTYSAREISENYGVDELLELQVQRAVGLARVDDPDVAVHMRADGEEAARQRFVEL 120

QY 121 GLNPDQVVLVVRVLAEGLSHAAEAMRYTALBAIMRPGATELIDIAKGSQALVSQIVPLLGP 180
DB 121 GLNPDQVVLVVRVLAEGLSHAAEAMRYTALBAIMRPGATELIDIAKGSQALVSQIVPLLGP 180

QY 181 MIQDMLFMQLRHMMETEAUNAGERAAGKPLPGARQVTVAFADLVGFTQLGWVSAEELGH 240
DB 181 MIQDMLFMQLRHMMETEAUNAGERAAGKPLPGARQVTVAFADLVGFTQLGWVSAEELGH 240

QY 241 LAGRLAGLARDLTAPPVWFITIGDAVNLVCPDPAPLLDTVLKLVVVDTDNFPRLRAG 300
DB 241 LAGRLAGLARDLTAPPVWFITIGDAVNLVCPDPAPLLDTVLKLVVVDTDNFPRLRAG 300

QY 301 VASGMVSRAGDWFPGSPVNVASRVTVGVARPGAVLVADSVREALGDAPADGFGQSFAGPR 360
DB 301 VASGMVSRAGDWFPGSPVNVASRVTVGVARPGAVLVADSVREALGDAPADGFGQSFAGPR 360

QY 361 RLIRGIRGVRLFRVRRGATRTGSGGAAQDDDLGSSP 397
DB 361 RLIRGIRGVRLFRVRRGATRTGSGGAAQDDDLGSSP 397

RESULT 3
QY3X03 PRELIMINARY; PRT; 385 AA.
AC QY3X03;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MAP2507c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017236; AAS04824.1; -.
DR GO; GO:0004383; F:guanylate cyclase activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001054; G_cyclase.
DR Pfam; PF00211; Guanylate_cyc; 1.
DR SMART; SM00044; CYCC; 1.
DR PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR Complete proteome; Lyase.
SQ SEQUENCE 385 AA; 41476 MW; A5A3203A0A1B2DD3 CRC64;

Query Match 73.7%; Score 1472.5; DB 2; Length 385;
Best Local Similarity 78.1%; Pred. No. 2.6e-85;
Matches 293; Conservative 29; Mismatches 50; Indels 3; Gaps 1;

```

QY 3 DHVREADDANIDLLGDLGGTARABRAKLVWLELQGITPPDEIRATNPPLLLATRHLLVGD 62
 DB 13 DDVHGDGDDRIEDLLDGLQGTARTARAEALVRWLLAQGITAEIRTNPPPLLLATRHLLIGD 72
 QY 63 DGTYSAREISNYGVDELLOQVORAVGLARVDDPDVAVHMRADGEAAARQRFVELGL 122
 DB 73 DGTYSVSTREISYGLDMALLQVORAVGLARVDDPDVAVHMRADGEAAARQRFVELGL 132
 QY 123 NPDQVVLVVRVLAEGLSHAAEAMRYTALPAIMRPGATELDDIAKGSQALVSQIVPLLGPMI 182
 DB 133 DFDQVVLVVAQLAEGLSRAAEVMRYALSALIMRPGATELEIAKASKALVTQIAPLLGPMI 192
 QY 183 QMFLFMQLHMMHMETEAVNAGERAAKPLPGARQVTVAFADLVGFTQGLGEVVAEELGHIA 242
 DB 193 QMFLFMQLHMMHMETEAVNAGERAAKPLPGARQITVAFADLVGFTQGLGEVVAEELGHIA 252
 QY 243 GRLAGLARDLTAPPVWFIKTIGDAMVLCVCPDPAIPLDVLKLVVVDDTNNFRLRAGVA 302
 DB 253 NRLAILARDVTVPPVRFVKTIGDAMVLCVCPDPAIPLDVLKLVVVDDTNNFRLRAGVA 312
 QY 303 SCMAVSRAAGDFGSPVNVASRVTVARPGAVLVADSVREALGDPAEADGFQWFSFAGPRRL 362
 DB 313 SGTAVSRAAGDFGSPVNVASRVTVARPGAVLVADSVVMDVMDVGDNGE---FSGSPFAGARRL 369
 QY 363 RGIQDVLRLFRVRG 377
 DB 370 KGIKNEVKLFRVRG 384

RESULT 4

CYAA_STRCO
 ID CYAA STRCO STANDARD; PRT; 399 AA.
 AC Q9WXC3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Adenylate cyclase (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl
 DE cyclase).
 GN Name=cya; Synonyms=cyaA;
 OS Streptomyces griseus.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1911;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9930325; PubMed=10376032;
 RA Kang D.K., Li X.M., Ochi K., Horinouchi S.;
 RT "Possible involvement of cAMP in aerial mycelium formation and
 secondary metabolism in Streptomyces griseus.";
 RL Microbiology 145:1161-1172(1999).
 CC -!- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
 CC -!- COPACTOR: Binds 1 magnesium ion per subunit (By similarity).
 CC -!- SIMILARITY: Belongs to the adenylyl cyclase class-3 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB018557; BAA76599.1; -;
 DR InterPro: IPR001054; G_cyclase.
 DR Pfam; PF00211; Guanylate_cyc; 1.
 DR SMART; SM00044; CYCC; 1.
 DR PROSITE; PS0125; GUANYLATE_CYCLASES_2; 1.
 KW CAMP biosynthesis; Lyase; Magnesium; Metal-binding.
 FT METAL 203 Magnesium (By similarity).
 FT METAL 247 Magnesium (By similarity).
 SQ SEQUENCE 399 AA; 43163 MW; C38C6A015F9E7685 CRC64;
 Query Match 16.3%; Score 326; DB 1; Length 399;

Best Local Similarity 31.1%; Pred. No. 1e-12;
 Matches 100; Conservative 47; Mismatches 141; Indels 34; Gaps 10;
 QY 45 IRATNPPLLLATRHLLV-GDDGTYSAREISENYGVDELLOQVORAVGLARVDDPDVAVH 103
 DB 32 VEPTDDPLAIRLEALILGADRRY-TFFQAARTAGVSMDLASFPWRAMAFADIGQAKALT- 89
 QY 104 MRADGEAAARQRFVELG-LNPDDQVVLVVRVLAEGLSHAAEAMRYTALPAIM---RPGAT 159
 DB 90 -EADVLALRLRLLAGLVEAGLLSEPMAIQVARSTGQTARLAENQIDSFLEGLTEPPEPGMT 148
 QY 160 ELDDIAKGSQALVSQIVPLLGPMIQMDLPMQLHMMHMETEAVNAGERAAKPLPGA----- 213
 DB 149 RTEV-----TYPILVELLLPELQBFVYVWRRLAA-----ATGRVQVQAADDEEMV 193
 QY 214 -RQVTVAFADLVGFTQGLGEVVAEELGHLAGRLAGLARDLTAP-PWRFIKTIGDAMVLCV 271
 DB 194 DRPRVRADLVGFTRLTRLEEEELGELVESFETTAADLVAPTAAGLVKTLGDELVLFAA 253
 QY 272 PDPAIPLDVLKLVVVDDTNNFRLRAGVAGSMVSRAGDFGSPVNVASRVTVARPG 331
 DB 254 DDAGTAAEIALRLIEAMSQDETMPALRVGIAFGTVTRMGDVGFTTVNLASRLTSTAPKD 313
 QY 332 AVLVDVADSVREAL---GDAPEAD 350
 DB 314 AVLVDGAFAPAKELVRHGEAPESE 335
 RESULT 5
 CYAA_STRCO
 ID CYAA STRCO STANDARD; PRT; 381 AA.
 AC P40135;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Adenylate cyclase (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl
 DE cyclase).
 GN Name=cya; OrderedLocusNames=SCO4928; ORFNames=SK13.20;
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94109670; PubMed=8282183; DOI=10.1016/0378-1097(93)90511-Y;
 RA Danchin A., Pidoux J., Krin E., Thompson C.J., Ullmann A.;
 RT "The adenylate cyclase catalytic domain of Streptomyces coelicolor is
 RT carboxy-terminal.";
 RL FEMS Microbiol. Lett. 114:145-152(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphoy L.D., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2)";
 RL Nature 417:141-147(2002).
 CC -!- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
 CC -!- COPACTOR: Binds 1 magnesium ion per subunit (By similarity).
 CC -!- SIMILARITY: Belongs to the adenylyl cyclase class-3 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----


```

CC CAMP.
CC -!- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
CC -!- COFACTOR: Binds 1 magnesium ion per subunit (By similarity).
CC -!- ENZYME REGULATION: Pyruvate-stimulated.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the adenyllyl cyclase class-3 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X57541; CAA40760.1; -.
CC PIR; S15273; OYFKQ.
CC InterPro; IPR001054; G_cyclase.
CC Pfam; PF00211; Guanylate_cyc; 1.
CC SMART; SM00044; CYCC; 1.
CC TIGRFAMs; TIGR01712; phage_NGA_met; 1.
CC PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
CC CAMP biosynthesis; Lyase; Magnesium; Metal-binding.
CC FT DOMAIN 31 60 Pyruvate binding (Potential).
CC FT DOMAIN 337 345 Catalysis and substrate binding
CC FT METAL 243 243 Magnesium (By similarity).
CC FT METAL 287 287 Magnesium (By similarity).
CC SQ SEQUENCE 403 AA; 43898 MW; 30293FF790E9AAE19 CRC64;
CC -----
Query Match 13.6%; Score 271; DB 1; Length 403;
Best Local Similarity 26.3%; Pred. No. 3.2e-09;
Matches 106; Conservative 64; Mismatches 181; Indels 52; Gaps 15;
CC
CC 4 HVREADDANTDGLGGTARAEAKLVFWLLE-QGITPDEIRATNPPLLATRLHVG 62
CC 27 HAPAAATPAESDPIL-----ELAEAMEGPIRIPAHTPEAVRDVTSAL---EKRLIG 74
CC 63 DGTYSAREISENYGVDELLELQVRAGVLARVDDPDVAVHMEADGEAAARQRFVELGL 122
CC 75 QREF-RRREVASAGVSLHSARKLWRAIGFPELSD-DEVFFTEADKKALGTWGMVREGA 132
CC 123 NPQOVVLVVRVLAEGLSHAAEMRYTALEAIMEPGATELDA-KGSQALVSQIVPLLPGM 181
CC 133 LTBETAI---SLMRSGQMTDRMVVMQIEALVEDMIANQLNSDRQARRQLFSLLPETIPA 189
CC 182 IQDMLFQOLRHMETEA-----VNAGERAAGKPLPGARQVTVAPADLVGF 226
CC 190 IEDLLYSWRQNSAVHRMALRVETGVAAYNODRGEDGGTLPPLAR--AVGFADLVSY 247
CC 227 TQGEVVVSABELGLAGRL-AGLARDLTAPPVWFVFIKTIGDVMVLCVCPDPAFLDITVLKLV 285
CC 248 TSLSRNRNERTLAQLVQREFAKCAEILISVGGRLVKTIGDEVLYVETPQAGNIALSL 307
CC 286 EVVDTNNPRLRAGVAGMVASRAGDFGSPNVASRVTVGVARPGAVL----VADSVRE 341
CC 308 RELAKDELFPQTRGAVGVWGRLLSRLGDIYGPVTVMMAARLTSLAEPGTVLDTAITNLR- 366
CC
CC 342 ALGDAPEADGQFQSFAGPRRLRGIRGDVRLFRVRGRATRTGSG 384
CC 367 --NDA----RFLVTAQEIATAVRGF-GDIQPIYELASG---EGAG 399
CC -----
CC
CC RESULT 8
CC YMI2_MYCTU STANDARD; PRT; 378 AA.
CC AC P64265; Q10400;
CC DT 01-OCT-1996 (Rel. 34, Created)
CC DT 25-OCT-2004 (Rel. 45, Last sequence update)
CC DT 25-JAN-2005 (Rel. 46, Last annotation update)
CC DE Hypothetical protein Rv2212/MT2268.
CC GN OrderedLocusNames=Rv2212, MT2268; ORFNames=MTCY190.23;

```

```

OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagsis K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.13.5479-5490.2002;
RA Frieschmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490 (2002).
CC -!- SIMILARITY: Belongs to the adenyllyl cyclase class-4/guanylyl
CC cyclase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; BX842579; CAA94255.1; -.
CC PIR; AE000516; AAK46554.1; ALT_INIT.
CC HSP; P19786; E70786.
CC TIGR; MT2268; -.
CC Tuberculin; RV2212; -.
CC InterPro; IPR009056; Cytochrome c.
CC InterPro; IPR001054; G_cyclase_.
CC Pfam; PF00211; Guanylate_cyc; 1.
CC SMART; SM00044; CYCC; 1.
CC PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
CC Complete proteome; Hypothetical protein.
CC FT DOMAIN 208 317 Guanylate cyclase.
CC SEQUENCE 378 AA; 39775 MW; AF26A9CFC13C91DC CRC64;
CC -----
Query Match 12.3%; Score 245; DB 1; Length 378;
Best Local Similarity 27.1%; Pred. No. 1.3e-07;
Matches 109; Conservative 61; Mismatches 152; Indels 80; Gaps 18;
CC
CC 9 DDANIDDLGLGGTAR-AERAKLVFWLLEQGITPDE-IRA-----TNP 50
CC 3 DSLDFDAL--EAAGIANPRERAGLTYLDELFTVEEMVQARRGLFCGLAGVLLWSGP 60
CC 51 PL-LLATRLHVGDDGTYSAREISENYGVDELLELQVRAGVLARVDDPDVAVHMEADGE 109
CC 61 PIYTLAT-----AADELGLSADDDVARAWSLGLT-VAGPDVPTLSQADV 104
CC 110 AAARQRFVELGLNPDQVVLVVRVLAEGLSHAAEMRYTALEAIMEPGATELDA-----TE 160

```

```
Db 105 ALATWVALKAL-VGEDGAFGLRLVLGTAMARLAEAE-----STMIRAGSPNIQMTHTHDE 158
QY 161 LDIAGSQAALVSIQIVPLLGPMLQDMFLMQLRHMMET-----EAYNAGERAAKGLPLPGARQ 215
Db 159 LATARAYRA-AAEFVPRIGALIDTV-----HRHHLASARTYFEGV-IGDTSA-----S 204
QY 216 VT--VAFADLVGFTQLGGEVSAEELHGLAGRLAGLARDLT-APPVWFIKTIGDAVMLVYCP 272
Db 205 VTCGIGFADLSSTFALTQALTPAQLODLLTFEDAATDVVHADGGRLVKFGIDAVVMWVSS 264
QY 273 DPAPLLDTVLKLVVVDTNNPRLRAGVAGSMVSRAGDWFSGSPVNVASRTVGVARPGA 332
Db 265 SPERLVRAAVDLVDHPGARAAELQVRAGLAYGTVLALNGDYFGNPNVLAARLVAAAAPGQ 324
QY 333 VLVADSVREALGDPAEADGFQWSPAGPRRLGIRGVDVRLFRV 374
Db 325 ILAAQLRDMPLDPWPAL-----AHGPLTLKGFDPVMAFEL 360

RESULT 9
YM35 MYCBO STANDARD; PRT; 378 AA.
AC P64266; Q10400;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Hypothetical protein MB2235.
GN OrderedLocusNames=MB2235;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12789972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duchoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -!- SIMILARITY: Belongs to the adenyl cyclase class-4/guanylyl
CC cyclase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; BX248341; CAD97088.1; -.
CC HSSP; P19754; 1ANK.
CC InterPro; IPR009056; Cytochrome c.
CC InterPro; IPR001054; G_cyclase_.
CC Pfam; PF00211; Guanylate_cyc; 1.
CC SMART; SM00044; CYCG; 1.
CC PROSITE; PS50125; GUANYLATE CYCLASES.2; 1.
CC Complete proteome; Hypothetical protein.
FT DOMAIN 208 317 Guanylate cyclase.
SQ SEQUENCE 378 AA; 39775 MW; AF26ASCFGL3C91DC CRC64;

Query Match 12.3%; Score 245; DB 1; Length 378;
Best Local Similarity 27.1%; Pred. No. 1.3e-07;
Matches 109; Conservative 61; Mismatches 152; Indels 80; Gaps 18;

QY 9 DDANIDLLGLDGGTAR-ABRAKLVEMLEGGITPDE-IRA-----TNP 50
Db 3 DLSDFDAL--BAAGIANPRERAGLTLYDELGFTVEEMVQAEGRRLFGLAGDVLWMSGP 60
QY 51 PL-LLATRLHVGDDGTYYVSAREISENYGVDLELLQVRQAVGLARVDDPDVAVVHRADGE 109
```

```
Db 61 PIYTLAT-----AADELGLSADDVARAWSLLGLT-VAGPDVPTLSQADVD 104
QY 110 AAAARAQRFVELGLNPDQVVLVVRVLAEGLSHAABAMRYTALEAIMRPGA-----TE 160
Db 105 ALATWVALKAL-VGEDGAFGLRLVLGTAMARLAEAE-----STMIRAGSPNIQMTHTHDE 158
QY 161 LDIAGSQAALVSIQIVPLLGPMLQDMFLMQLRHMMET-----EAYNAGERAAKGLPLPGARQ 215
Db 159 LATARAYRA-AAEFVPRIGALIDTV-----HRHHLASARTYFEGV-IGDTSA-----S 204
QY 216 VT--VAFADLVGFTQLGGEVSAEELHGLAGRLAGLARDLT-APPVWFIKTIGDAVMLVYCP 272
Db 205 VTCGIGFADLSSTFALTQALTPAQLODLLTFEDAATDVVHADGGRLVKFGIDAVVMWVSS 264
QY 273 DPAPLLDTVLKLVVVDTNNPRLRAGVAGSMVSRAGDWFSGSPVNVASRTVGVARPGA 332
Db 265 SPERLVRAAVDLVDHPGARAAELQVRAGLAYGTVLALNGDYFGNPNVLAARLVAAAAPGQ 324
QY 333 VLVADSVREALGDPAEADGFQWSPAGPRRLGIRGVDVRLFRV 374
Db 325 ILAAQLRDMPLDPWPAL-----AHGPLTLKGFDPVMAFEL 360

RESULT 10
Q9CC18 PRELIMINARY; PRT; 324 AA.
ID Q9CC18;
AC Q9CC18;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein ML1399.
GN OrderedLocusNames=ML1399;
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002; DOI=10.1038/35059006;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C.M., Harris D.E.,
RA Mungall K.L., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L.D., Oliver K., Quail M.A., Rajadream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus."
RL Nature 409:1007-1011(2001).
DR EMBL; AL583921; CAC31780.1; -.
DR PIR; A87084; A87084.
DR Leproma; ML1399; -.
DR GO; GO:0004383; F:guanylate cyclase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR01054; G_cyclase.
DR InterPro; IPR010916; TONB_Box N.
DR Pfam; PF00211; Guanylate_cyc; 1.
DR SMART; SM00044; CYCG; 1.
DR PROSITE; PS50125; GUANYLATE CYCLASES.2; 1.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
KW Complete proteome; Hypothetical protein; Lyase.
SQ SEQUENCE 324 AA; 34670 MW; 5DA2095EE0186AGA CRC64;

Query Match 9.5%; Score 190.5; DB 2; Length 324;
Best Local Similarity 25.9%; Pred. No. 0.00031;
Matches 95; Conservative 46; Mismatches 139; Indels 87; Gaps 15;

QY 42 PDEIRATNPPLLLATRLHVLVGDDGTYYVSAREISENYGVDLELLQVRQAVGLARVDDPDVAV 101
```



```
Db 28 PDEDTVTVPAPALIRSSVS-----MENAQAQLHTTNRSPRLVAMVR----- 69
QY 102 VHMRADEGAARARQRFVGLNLPDQVVLVVRVLAEGLSHAAEAMRYTALEAIRMPOATEL 161
Db 70 -----RARRL--LFGDPLD-----FGDPLS-----TAGEGGPRAAARAA 100
QY 162 DIAKGSQALVSQIVLPLGPMIQDMLFMQLRHMMHMETEAVNAGERAACKPLPGARQVTVAF 221
Db 101 DRLLGDRGAASRVLSVLQVQAL-----TEAI-----ARRPV--NPEVTLVFT 143
QY 222 DLVGFQTLGGEVSAEELHGLHAGRLAGLARDLTAPPV-----WFIKTTIGDAMVLVCPDPAPL 277
Db 144 DLVGFS--GWSLQAGDEATLA--LLRQVARAVESPLLDAGCHIYKRMGDDGIMAVFRDPSVA 200
QY 278 LDTVLKLVV-----VDTDNPNPRLRAGVAGSMVSRAGDWFSGSPVNVASRVTVGVARPGAVL 334
Db 201 VOAVLAATEAMKSVVEGGYTPRIRVGIHTGRPORLAADWLGVVDVNTAARVMERATKGGIM 260
QY 335 VADSVREALGDPAEDGFQWSFAGPRRLR-----GIRGDVRLFRVRRGATRTGSGG 385
Db 261 ISGPTLUDL-----POSOLKELGII--TRRVKPMPTSKFTGLIPDMVVIYRIKARRELTA 316
QY 386 AAQDDDL 392
Db 317 TAQTNSL 323

RESULT 11
Q989E3
ID Q989E3 PRELIMINARY; PRT; 481 AA.
AC Q989E3
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE M16463 protein.
GN OrderedlocusNames=m16463;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RA "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003009; BAB52754.1; -.
DR GO; GO:0004383; F:guanylate cyclase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001054; G:cyclase.
DR Pfam; PF00211; Guanylate_cyc; 1.
DR SMART; SM00044; CYCC; 1.
DR PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR Complete proteome; Lyase.
KW SEQUENCE 481 AA; 52579 MW; 5BFD590AE53A061 CRC64;

Query Match 9.3%; Score 186.5; DB 2; Length 481;
Best Local Similarity 26.4%; Pred. No. 0.00087;
Matches 74; Conservative 42; Mismatches 93; Indels 71; Gaps 12;

QY 143 EMARYTALEAIRMPOATELDIA-----KGSQALVS---QIVPLL-----GPM 181
Db 226 DGRKFLPAMPVPPGPIVIEVANRGPVRSLLVINWNPPELVALTPTKLTDFDPVYSGGAL 285
QY 182 IQDMLFMQLRHMMHMETEAVNAGERAACKPLPGARQVTVAFADLVGTLQGEVSAEELG-- 239
```

```
Db 286 LARQTFRQLFR-----SERVDEKESGLGIQITFLFTDLKSGTAM-----YERLGD 331
QY 240 -----HLARLGLAGLARDLTAPPVWFIKTTIGDAMVLVCPDPAPLDTVLKLVVVD 290
Db 332 NAYALVREHFALVNAVQOHSQA---VVKTTIGDAMVAVFSQPSDAISAALHVFEEIDRF 387
QY 291 --DNNFPRL-----RAGVASCMAVSRAGDWFSGSPVNVASRVTVGVARPGAVLVADSVREA 342
Db 388 NGDHGPGIILKIGAHCGPSIAVTLDNLDYFGQTVNVAAARVQSLAEAGQICLS-----EA 443
QY 343 LGDAPEADGFQWSFAGPR-----RLRGIRGDVRLFRVR 376
Db 444 LYSAP---GVREMLAGHRVVAFDAPLRGVQGDATVVRIMR 480

RESULT 12
Q7MWV9
ID Q7MWV9 PRELIMINARY; PRT; 412 AA.
AC Q7MWV9
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Putative regulatory protein.
GN Name=tutC; ORFNames=PHG383;
OS Alcaligenes eutrophus (Ralstonia eutropha).
OG Plasmid megaplasmid pHG1.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Wautersia.
OX NCBI_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H16;
RX MEDLINE=22830147; PubMed=12949488; DOI=10.1016/S0022-2836(03)00894-5;
RA Schwartz E., Henne A., Cramm R., Eitinger T., Friedrich B.,
RA Gottschalk G.;
RT "Complete Nucleotide Sequence of pHG1: A Ralstonia eutropha H16
RT Megaplasmid Encoding Key Enzymes of H2-based Lithoautotrophy and
RT Anaerobiosis.";
RL J. Mol. Biol. 332:369-383(2003).
DR EMBL; AY305378; AAP86132.1; -.
DR HSSP; P26769; 1AB8.
DR GO; GO:0004383; F:guanylate cyclase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0000155; P:two-component sensor molecule activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR001054; G:cyclase.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS.
DR InterPro; IPR000700; PAS-assoc_C.
DR Pfam; PF00211; Guanylate_cyc; 1.
DR SMART; SM00044; CYCC; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 1.
DR TIGRfams; TIGR00229; sensory box; 1.
DR PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR PROSITE; PS50113; PAC; 1.
DR PROSITE; PS50112; PAS; 1.
KW Lyase; Plasmid.
SQ SEQUENCE 412 AA; 46131 MW; 3A1FDFF742F08BDA CRC64;

Query Match 8.6%; Score 172; DB 2; Length 412;
Best Local Similarity 25.9%; Pred. No. 0.006;
Matches 96; Conservative 50; Mismatches 117; Indels 108; Gaps 20;

QY 76 YGVDELQVRVORAVGLARVDP-----PDVV-----HMRDV 107
Db 52 HGTMLINRAALEAIGI-RMDDIRKPFWEARWVWVSPETQALLRELIIQASOGFEVRCD 110
QY 108 ----GEAARAQRVFEGLNP---DQVVLVVRVLAEG-----LSHAAEAMRYTALEA-- 152
Db 111 PENYQAECEEKIIVDFSLPIRDSNGNVVFLPEGRNITEKKRAEALAKNEELQASW 170
```

```

QY 153 -IMRPGATELDIAKGSQALVSQIVPLGLPMIQDMLFMOLRHHMETEAVNAGER--AAGKP 209
Db 171 ELIRQRDEL-----QSLYDRI--LAQKLSERLLNLNLLPYPAERLUKARPDLIASGVP 222
QY 210 ---LPGARQVTVAFADLVGFTQLGGEVVAEELGHLIAGRL-----AGLARDLTAPP 256
Db 223 ELIADSPFEVTVLFPADIVAFTRFSAGMSPEQLVAITNEIFTEFDFTIAEVRGLEK----- 276
QY 257 VWFIKTIGDANMLVC--PDAPLLDTVLKLVVV-----DTNNFP-----RLRAGVAS 303
Db 277 ---IKTIGDAYMAAAGLEPAP--DHAVRAAHMALNMSDALMHFNQHTGFNLQMRIGINS 331
QY 304 GWAVSRAG-----DWFQSPVNVASRVTVARPGAVLVADSVREALGDAPADGQFOW 354
Db 332 GAVV--AGVIGKRFYDLWGAVNIAISMESHGVAGRVQVTDATRLLGES-----FQF 384
QY 355 SFAGPRLRLGI 365
Db 385 EERGSIEAKGI 395

RESULT 13
Q740J3
ID Q740J3 PRELIMINARY; PRT; 322 AA.
AC Q740J3
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DE Hypothetical protein.
GN OrderedLocusNames=MAP1357;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K10.
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
RE EMBL; AE017231; AAC03674.1; -.
DR GO; GO:0004383; F:guanylate cyclase activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001054; G_cyclase.
DR Pfam; PF00211; Guanylate_cyc; 1.
DR SMART; SM00044; CYCC; 1.
DR PROSITE; PS00125; GUANYLATE_CYCLASES_2; 1.
KW Complete proteome; Lyase.
SQ SEQUENCE 322 AA; 34884 MW; C6FC4F3744409D80 CRC64;

Query Match 8.3%; Score 166.5; DB 2; Length 322;
Best Local Similarity 29.2%; Pred. No. 0.01; Mismatches 30; Indels 70; Gaps 7;
Matches 54; Conservative 30;

QY 211 PGARQTVAFADLVGFT---QLGEVSAEELGHLIAGRLAGLARDLTAPPV-----WPIK 261
Db 131 PANPEVTLVFTDLVGFTSLQAGDDAALLRLQVARAV-----EPPLDAGGHIVK 182
QY 262 TIGDANMLVCDDPAPLLDTVLKLVV---VDTNNFPRLRAGVAGMAVSRAQNWFGSPV 318
Db 183 RMGDGLMAVFGDPTVAVRAVLAALKALRSVEVAGVTPRMVGIHTGRPQLASDLWLGVDV 242
QY 319 NVASRVTVARPGAVLVADSV-----REALGDAPADGQFOWSFAGPRLRLGIRGQVR 370
Db 243 NIAARWMERATKGGIMVSSSTLDH1PQSELDALG--VEAKTRKPVFGPKP-AGMPADLA 299
QY 371 LFRVR 375
Db 300 IYRLK 304

RESULT 14
Q89SW2
ID Q89SW2 PRELIMINARY; PRT; 527 AA.

```

```

AC Q89SW2;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Two-component hybrid sensor and regulator.
GN OrderedLocusNames=blr2288;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobiium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA110.
RE MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idegawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RL DNA Reg. 9:189-197(2002).
DR EMBL; AP005943; BAC47553.1; -.
DR HSSP; P26769; IAZS.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004383; F:guanylate cyclase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0007600; P:sensory perception; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .; IEA.
DR InterPro; IPR011006; Chey_like.
DR InterPro; IPR001054; G_cyclase.
DR InterPro; IPR003661; His_kinA_N.
DR InterPro; IPR009082; His_kin_homodim.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF00211; Guanylate_cyc; 1.
DR Pfam; PF00512; Hiska; 2.
DR Pfam; PF00072; Response_reg; 1.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00044; CYCC; 1.
DR SMART; SM00388; Hiska; 2.
DR SMART; SM00448; REC; 1.
DR PROSITE; PS00125; GUANYLATE_CYCLASES_2; 1.
DR PROSITE; PS00110; RESPONSE_REGULATORY; 1.
KW Complete proteome; Lyase; Phosphorylation; Sensory transduction.
SQ SEQUENCE 527 AA; 58612 MW; 264BBBD4E0GACAF5 CRC64;

Query Match 8.2%; Score 163; DB 2; Length 527;
Best Local Similarity 26.9%; Pred. No. 0.03; Mismatches 47; Indels 102; Gaps 21;
Matches 104; Conservative 47;

QY 13 IDLLGLDGLGT---ARAERAKLVEWLEQGITPDEIRATNPPLLATRLHLVGDGTYVS- 68
Db 164 VSDLMSRSLGPRVQVAATETGSI-----LVVDDIEANRD---LLSRRLT-RGHRVSS 211
QY 69 -----ARETSENVGVDLELLQVRQVAVGLARVDDPDPAVVMRADGEAAAARQRPVELGL 122
Db 212 VAGGQALQALANDEFDLVLLDLN-----MPDINGLDLVVRMKAD-----ERLRRI- 257
QY 123 NPDQVVLVVRVLARGLSHAAMEYVTALEAMRRCATELIDIAKGSQALVSQIVPLLGPMI 182
Db 258 ----PVIMITALATES-AVRCIEAGAEVDLPKFPDPIILLRLARINACLUHKK-----RW 305
QY 183 QDMLFMQLRHHMETEA-----VNAGERAAGKPLPGARQVTVAFADLVG 225
Db 306 RDEQKYLRIIEETAKFERULLLTILPRQVIGRLNHGEMIAIDREFG---VSVLPADLVG 362
QY 226 FTQLGEVVS-----ABELGHLIAGRLARDLTAPPVWFIKTIGDANMLVC---PDAPLLD 279
Db 363 FTEHSSRVTPAAMVEYLNRLFSEFDALAREL---GVEXIKTIGDAYMAVAGLPDPNP--D 417

```

```
OY 280 TVLKLVEV-----VDTNNFP-RLRAGVAGMAVSRAG-----DWFGSPVN 319
DB 418 ATAIAIARMAIMIDIRGVNSHFGWPLQIRIGHSGPVV--AGIIGAHRFIYDVWMDTVN 475
OY 320 VASRVTGVARPGAVLVA-DSVREALG 344
DB 476 VASRLAAYSILPNRIHVSRDIARHLVG 501

RESULT 15
P94982
ID P94982 PRELIMINARY; PRT; 316 AA.
AC P94982;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=Rv1647;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekaiia F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., Mclean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
DR EMBL; BX842577; CAB06641.1; -.
DR PIR; B70620; B70620.
DR Tuberculist; Rv1647; -.
DR DR GO:0004383; F:guanylate cyclase activity; IEA.
DR DR GO:0016829; F:lyase activity; IEA.
DR DR GO:0007242; P:intracellular signaling cascade; IEA.
DR DR InterPro; IPR001054; G_cyclase.
DR DR SMART; SM00044; CYCC; 1.
DR DR PROSITE; PS0125; GUANYLATE CYCLASES 2; 1.
KW Complete proteome; Hypothetical protein; Lyase.
SQ SEQUENCE 316 AA; 33956 MW; 34E1AD48472DA31D CRC64;

Query Match 8.1%; Score 162; DB 2; Length 316;
Best Local Similarity 26.5%; Pred.No. 0.019;
Matches 62; Conservative 27; Mismatches 63; Indels 82; Gaps 9;

OY 196 TEAVNAGERAAGKPLPGARQVTVAFADLVGF-----TOLGEVVSAREL----- 238
DB 117 TEAVSR-----RPAPEVTLVFTDLVGFSTWLSHAGDADATLTLLRQVARAVESPLLD 168
OY 239 --CHLAGRLAGLADLTAPPVWFIKTIGDVMVLVCPDAPPLDQTLV---KLVEVVDTDNN 293
DB 169 AGGHI-----VKELGGGIMAVFRNPTVALRAVLVAQDAVKSLEVOGY 210
OY 294 FPLRLAGVAGMAVSRAGDWFSGSPVNVASRVTGVARPGAV-----LVADSVREALGD 345
DB 211 TPEMRIGIHTGRPQRLAADWLGVGVNIAARVMERATKGGIMISQPTLDLIPQSELDALG- 269
OY 346 APEADGFQWFSAGPRRLR-----GIRGDLFRVRRGATRTGSGGAQDD 390
DB 270 -----VVARRVRKPVFASKPTGIPPDIAIYRI-----KTVSESTAADN 307
```

Search completed: October 15, 2005, 04:35:55
Job time : 89.9255 secs

THIS PAGE BLANK (000710)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:20:35 ; Search time 118.566 Seconds
(without alignments)
1454.845 Million cell updates/sec

Title: US-10-617-038-8
Perfect score: 2332
Sequence: 1 MVPEPNLAGATCAEWIGRPP.....RLAVIAAKVITGRKLSRRPL 446

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A: Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2332	100.0	446	ADJ37287	M. tuberc
2	621	26.6	446	ABJ25493	Aspergill
3	621	26.6	468	ABJ25791	Aspergill
4	621	26.6	468	ABJ26391	Aspergill
5	621	26.6	468	ABJ26093	Aspergill
6	621	26.6	468	AD86052	Aspergill
7	501	21.5	441	ABJ76130	Zygoascus
8	491.5	21.1	460	ADP98845	C. albica
9	439	18.8	468	ADP73893	Candida a
10	438.5	18.8	388	AA870955	S. spinos
11	437.5	18.8	465	ABG74896	C. paraps
12	437.5	18.8	465	ADP08158	Lipase/ac
13	437.5	18.8	471	ABG74897	C. paraps
14	437.5	18.8	471	ADP08160	Lipase/ac
15	437.5	18.8	487	ADP33671	B. bassia
16	425.5	18.2	418	ADL27517	Amino aci
17	416.5	17.9	466	ABP73905	Candida a
18	383	16.4	462	AA878230	Candida a
19	383	16.4	462	AA848084	C. antart
20	383	16.4	463	AA848084	C. antart
21	335.5	14.4	320	AA8739309	SpmM prot
22	331.5	14.2	310	ABP57690	Saccharop
23	274.5	11.8	435	AA876628	Corynebac
24	274.5	11.8	435	AA879149	Corynebac
25	272.5	11.7	408	AA891382	C. glutami

26	181.5	7.8	402	6	ADA33693	Acinetoba
27	162	6.9	93	5	ABP09552	Human ORP
28	124	5.3	424	6	ADA33090	Acinetoba
29	120	5.1	631	4	ABG09020	Novel hum
30	117	5.0	390	8	ADN25118	Bacterial
31	116.5	5.0	753	5	ABU05595	M. tuberc
32	114.5	4.9	1827	4	AA881210	Mycobacte
33	112.5	4.8	1847	4	ABJ31557	Pimaricin
34	112.5	4.8	7068	4	AAE10142	Streptomy
35	112.5	4.8	9477	4	AAE10144	Streptomy
36	112	4.8	3724	2	AAW22608	Platenoli
37	112	4.8	3724	2	AAW23718	Platenoli
38	111	4.8	494	8	ADS28840	Bacterial
39	109	4.7	472	3	ABJ32526	Bacterial
40	109	4.7	472	7	ADE10298	S. lavend
41	107	4.6	489	6	ADA33347	Acinetoba
42	107	4.6	1436	2	AAW41361	Receptor
43	107	4.6	1436	7	ADB79775	Rat putac
44	106.5	4.6	510	8	ADJ49547	Oil-assoc
45	106.5	4.6	1402	4	AA881115	Mycobacte

ALIGNMENTS

RESULT 1
ADI37287
ID ADI37287 standard; protein; 446 AA.
XX
AC ADI37287;
XX
DT 22-APR-2004 (first entry)
XX
DE M. tuberculosis low oxygen induced antigen Rv1592c SEQ ID NO:8.
XX
KW mycobacterial infection; vaccine; tuberculosis;
KW Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;
KW low oxygen induced antigen.
XX
OS Mycobacterium tuberculosis.
XX
PN WO2004006952-A2.
XX
PD 22-JAN-2004.
XX
PF 08-JUL-2003; 2003WO-DK000477.
XX
PR 13-JUL-2002; 2002DK-00001098.
XX
PA (STAT-) STATENS SERUM INST.
XX
PI Andersen P, Rosenkrands I, Stryhn A;
XX
DR WPI: 2004-122778/12.
DR N-PSDB; ADI37332.
XX
PS Claim 3; SEQ ID NO 8; 76pp; English.
XX
PT Use of one or more polypeptides or their fragments, which are expressed during the latent stage of the mycobacterial infection, and/or nucleic acids encoding the polypeptides, for a therapeutic vaccine against tuberculosis.
XX
CC The present invention describes polypeptides or their fragments, which are expressed during the latent stage of a mycobacterial infection, and/or nucleic acids encoding the polypeptides, which are useful for creating a therapeutic vaccine against tuberculosis. Also described: (1) a therapeutic vaccine against tuberculosis comprising one or more polypeptides; (2) a method for treating an animal, including a human being, with tuberculosis caused by virulent mycobacteria, e.g. by Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for immunising an animal, including a human being, against tuberculosis caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis

caused by virulent mycobacteria in an animal, including a human being; (5) a method for diagnosing previous or ongoing infection with a virulent mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis infection in a subject. The polypeptides have antibacterial activities, and can be used in vaccines and in gene therapy. The polypeptides are useful for the manufacture of a therapeutic vaccine for treating an individual who is infected by a virulent mycobacterium, e.g. M. tuberculosis, and who is not vaccinated with BCG against tuberculosis. The present sequence represents a low oxygen induced antigen, which is used in the exemplification of the present invention.

Sequence 446 AA:

Query Match 100.0%; Score 2332; DB 8; Length 446;

Best Local Similarity	100.0%;	Pred. No. 2e-202;	Indels	0;
Matches	446;	Conservative	0;	Mismatches
Matches	446;	Conservative	0;	Gaps

Qy	1	MVEPGNLAGATGAEWIGRPPHEELQRKVRPILLPSDDPFYFPFAGYQHAVPGTVLRSDRVE	60
Db	1	MVEFGNLAGATGAEMIGRPPHEELQRKVRPILLPSDDPFYFPFAGYQHAVPGTVLRSDRVE	60
Qy	61	LAFMGILIPOPVTATOLLRYRTNNMGNBPATVTTIVPAELAPGGTCPLLSYQCAIDAMSS	120
Db	61	LAFMGILIPOPVTATOLLRYRTNNMGNBPATVTTIVPAELAPGGTCPLLSYQCAIDAMSS	120
Qy	121	RCFPSVALRRRAKALGSLTQMELLMI SAALAEGWAVSVDPDHEGPKGLWGSPYPBGPYRLVD	180
Db	121	RCFPSVALRRRAKALGSLTQMELLMI SAALAEGWAVSVDPDHEGPKGLWGSPYPBGPYRLVD	180
Qy	181	GIRAALNSERVGLSPATPIGLWGYSGGGISAWAAEACGEYAPDLDTVGAVLGSVPBGDLG	240
Db	181	GIRAALNSERVGLSPATPIGLWGYSGGGISAWAAEACGEYAPDLDTVGAVLGSVPBGDLG	240
Qy	241	HTFRRLNGTLTAGLPALVVAALQHSYPGLARVIKEHANDERGRQLLEOTBMTTVDVAIRM	300
Db	241	HTFRRLNGTLTAGLPALVVAALQHSYPGLARVIKEHANDERGRQLLEOTBMTTVDVAIRM	300
Qy	301	AGRDMDGFLOEPLEDILSTPEISHVFGBTKLGSAVPTPPVLIVQAVHDYLLDVSDIDALA	360
Db	301	AGRDMDGFLOEPLEDILSTPEISHVFGBTKLGSAVPTPPVLIVQAVHDYLLDVSDIDALA	360
Qy	361	DSVTAGGANVTYHRDLFSBHVSLHPLSAPMTLLRWLTDRFPAGKPULTDHRVRTTTPTFINPM	420
Db	361	DSVTAGGANVTYHRDLFSBHVSLHPLSAPMTLLRWLTDRFPAGKPULTDHRVRTTTPTFINPM	420
Qy	421	TYAGMARLAVIAAKVITGRKLSRRRL	446
Db	421	TYAGMARLAVIAAKVITGRKLSRRRL	446

RESULT 2

RESUL 2
ABJ25493
ID ABJ25493 standard: protein: 446 AA.

AC AB.T25493:

DT 16-APR-2003 (first entry)

XX DE Aspergillus fumigatus essential gene protein #151.

XX Fungicide; cytostatic; essential gene; *Aspergillus fumigatus*; infection;
KW cancer; contamination; biofilm; antibody; immune response.
KW

XX OS Aspergillus fumigatus.

XX PN WO200286090-A2

31-OCT-2002

23-APR-2002: 2002WO-US013142.

XX
PR 23-APR-2001; 2001US-0285697P.

Db 49 VRPL---EDPFYAPKGFESTVPGTILRWPNPNPISAFGAPINLAASYQLLYRSTDSF 105
Qy 85 GNPEATVTVIIPAEAPGQTCPLLSYQCAIDAMSSRCFSPVALRRRAKALGSLT-----LDSATDDELG 157
Db 106 GEPIAAASTILVPHN---ADNTKLLSFQAAEDAAHPNCAPSFAFQ-----LDSATDDELG 157
Qy 140 -----QWELLMISAALAEAGWAVSPDHEGPKLWGSYPYEGYRVLDGIRAAALNSERVGLSP 195
Db 158 LIMPOAELVLIITAAALDKGWWTVDPDLGNATFLANNLSGHVVLDNIRALRSSAFSGISP 217
Qy 196 ATPIGLWYSGGLASAWAAEACGEYAPDLIDIVGAVLGSFVPGDLGHTFRLNGTLLAGLP 255
Db 218 KATITLWYSGGLASGLAAELRASYAPELNIAAGALGGTVPKIMPVNTVNGKIYAGLL 277
Qy 256 ALVVAALOHSPGLARVKEHANEGRQLEQTEMTTVDVIRWAGRDMGDFLDEPLED 315
Db 278 PAGMOGLSNEYPAIEKILYDHLVPAKKADPVTKNLKLCIVEDLLTYSFQDFYRITD--AN 335
Qy 316 ILSTPEISHVFGDTKLGSAPVTPPVLIQVAVHDYLDVSDIDALADSYTAGGANVTYHRD 375
Db 336 MLKDPVTRVLGENAMGQHVDPIDPLFVYKSTNDVSPVGDTDALVSGYCAAGKVEYYRD 395
Qy 376 LFESEHVSLLHPLSAPMTLRLWTRFACKPL-TDHRVTRTWTPTTFNPMYAGM-----ARL 428
Db 396 ELNSHATMAVIGVGNALLWLKDRMNGVPARAGCKTQTALTGLDPRTLAVGLIDLIKVL 455
Qy 429 AVIAAKVIT 437
Db 456 ALLSAPVGT 464

RESULT 6

AD86052
ID ADR86052 standard; protein; 468 AA.
XX ADR86052;
AC ADR86052;
XX 04-NOV-2004 (first entry)
DT 04-NOV-2004 (first entry)
XX Aspergillus fumigatus essential gene protein #102.
DE Fungicide; Aspergillus fumigatus infection; Farmer's lung disease;
KW drug screening.
XX Aspergillus fumigatus.
XX WO2004067709-A2.
XX 12-AUG-2004.
XX 16-JAN-2004; 2004WO-US001099.
XX 17-JAN-2003; 2003US-0441281P.
PR 13-JUN-2003; 2003US-0478196P.
XX (ELIT-) ELITRA PHARM INC.
PA (ELIT-) ELITRA CANADA LTD.
XX Jiang B, Hu W, Lemieux S, Roemer T;
XX WPI; 2004-594200/57.
DR N-PSDB; ADR85465.
XX New purified or isolated Aspergillus fumigatus nucleic acid molecule
PT encoding a gene product, useful for diagnosing and/or treating invasive
PT fungal infections, such as farmer's lung disease.
XX Claim 1; SEQ ID NO 3102; 164pp; English.
PS The present invention relates to Aspergillus fumigatus genes that are
CC essential and are potential targets for drug screening. The methods and
CC compositions of the present invention are useful for diagnosing and/or

CC treating invasive Aspergillus fumigatus infection, including the allergic
CC forms of the disease, such as Farmer's lung disease. They can also be
CC used in various drug discovery purposes, such as expression of the
CC recombinant protein, hybridization assay and construction of nucleic acid
CC arrays. The present sequence represents an Aspergillus fumigatus
CC essential gene protein sequence, used during diagnosis and drug
CC development in the invention. These genes share a high degree of sequence
CC conservation with known essential genes of candida albicans. The sequence
CC data for this patent is not represented in the printed specification, but
CC was obtained in electronic format from WIPO.
XX
SQ Sequence 468 AA;
Query Match 26.6%; Score 621; DB 8; Length 469;
Best Local Similarity 35.0%; Pred. No. 3.5e-47;
Matches 150; Conservative 74; Mismatches 173; Indels 32; Gaps 9;
Qy 28 VRPLPSDDPFYFAGYQHAVPGTVLRSRDV--ELAFMGLIPQPVAT-QLLYRTNMY 84
Db 49 VRPL---EDPFYAPKGFESTVPGTILRWPNPNPISAFGAPINLAASYQLLYRSTDSF 105
Qy 85 GNPEATVTVIIPAEAPGQTCPLLSYQCAIDAMSSRCFSPVALRRRAKALGSLT-----LDSATDDELG 157
Db 106 GEPIAAASTILVPHN---ADNTKLLSFQAAEDAAHPNCAPSFAFQ-----LDSATDDELG 157
Qy 140 -----QWELLMISAALAEAGWAVSPDHEGPKLWGSYPYEGYRVLDGIRAAALNSERVGLSP 195
Db 158 LIMPOAELVLIITAAALDKGWWTVDPDLGNATFLANNLSGHVVLDNIRALRSSAFSGISP 217
Qy 196 ATPIGLWYSGGLASAWAAEACGEYAPDLIDIVGAVLGSFVPGDLGHTFRLNGTLLAGLP 255
Db 218 KATITLWYSGGLASGLAAELRASYAPELNIAAGALGGTVPKIMPVNTVNGKIYAGLL 277
Qy 256 ALVVAALOHSPGLARVKEHANEGRQLEQTEMTTVDVIRWAGRDMGDFLDEPLED 315
Db 278 PAGMOGLSNEYPAIEKILYDHLVPAKKADPVTKNLKLCIVEDLLTYSFQDFYRITD--AN 335
Qy 316 ILSTPEISHVFGDTKLGSAPVTPPVLIQVAVHDYLDVSDIDALADSYTAGGANVTYHRD 375
Db 336 MLKDPVTRVLGENAMGQHVDPIDPLFVYKSTNDVSPVGDTDALVSGYCAAGKVEYYRD 395
Qy 376 LFESEHVSLLHPLSAPMTLRLWTRFACKPL-TDHRVTRTWTPTTFNPMYAGM-----ARL 428
Db 396 ELNSHATMAVIGVGNALLWLKDRMNGVPARAGCKTQTALTGLDPRTLAVGLIDLIKVL 455
Qy 429 AVIAAKVIT 437
Db 456 ALLSAPVGT 464

RESULT 7

ABB76130
ID ABB76130 standard; protein; 441 AA.
XX ABB76130;
AC ABB76130;
XX 15-JUL-2002 (first entry)
DT Zygoscus hellenicus phospholipase.
XX Phospholipase; enzyme; baking; detergent.
KW Zygoscus hellenicus.
XX Zygoscus hellenicus.
XX Key Location/Qualifiers
FT Peptide 1..17 /label= Signal_peptide
FT Protein 18..441 /label= Mature protein
FT /note= "the mature protein is specifically claimed in
XX Claim 1"
PN WO200224881-A1.

```
XX PD 28-MAR-2002.
XX PF 18-SEP-2001; 2001WO-DK000600.
XX PR 25-SEP-2000; 2000DK-00001411.
XX PR 02-OCT-2000; 2000US-0237175P.
XX PA (NOVO ) NOVOZYMES AS.
XX PI Nielsen TAB, Ishii M, Patkar SA, Teutsami N;
XX DR WPI; 2002-383187/41.
XX DR N-PSDB; ARL56980.
XX PT New phospholipase from Zygoascus hellenicus useful for preparing a dough
XX PT or baked product, for reducing content of phosphorus in vegetable oil,
XX PT for production of animal feed and for partial hydrolysis of
XX PT phospholipids.
XX PS Claim 1; Page 17-19; 24pp; English.
XX CC The present sequence is the protein sequence of a novel phospholipase
XX CC from Zygoascus hellenicus strain CBS 4075. The enzyme has phospholipase A
XX CC activity and is able to hydrolyse lecithin, releasing fatty acid. It has
XX CC no lipase activity and shows no activity towards tributyrin. The
XX CC isoelectric point is about 4.2. The invention provides claimed
XX CC phospholipase polypeptides and nucleic acid sequences, expression
XX CC vectors, recombinant host cells, and a method for producing the
XX CC phospholipase by cultivating the host cells. A dough composition
XX CC comprising the phospholipase, a detergent composition comprising a
XX CC surfactant and the phospholipase, and a process for reducing the
XX CC phosphorus content of a vegetable oil using the phospholipase are
XX CC claimed. The phospholipase is also useful for: improving the
XX CC filterability of an aqueous solution or slurry of carbohydrate origin;
XX CC partial hydrolysis of phospholipids, particularly lecithin, to obtain
XX CC improved phospholipid emulsifiers; production of animal feed; and for
XX CC reducing the content of phospholipid in an edible oil
XX CC
XX SQ Sequence 441 AA;

Query Match      21.5%; Score 501; DB 5; Length 441;
Best Local Similarity 33.9%; Pred. No. 2.5e-36;
Matches 128; Conservative 62; Mismatches 166; Indels 22; Gaps 7;

QY 30 PLLSDDDFPPAGYQHAVGTULSRDVL--ELAFWGLIQQPVAT--QLLYRTNNMYGN 86
DB 20 FVAISDDPYNAPANLDSLADGTIIRSETPTNSLAALVLPQNVQSKYLLYKSTDDLNR 79
QY 87 PEATVTIVPAELAPGQTCPLLSYQCAIDAMSSRCFPPSYALRRRAKALSLTQWELMI 146
DB 80 TFATATTIIVPN---ADYTKLVSFQIAEDSAYNCAPSYALQFASDPDALLSQVEFVL 136
QY 147 SAALAEGWASVPDHEGKGLMGSPYEGYRVLDGIRAAALNSERY-GLSPATPIGLMGYS 205
DB 137 SAMLGLGVVNVDPYEGFRSAFTVQVAGRAVLDSLRAALASSSTIGTINSKARIALMGYS 196
QY 206 GGLIASAWAECAGYAPDLIDIVGAVLSPGVGDGLHTFRRRLNGTLLAGLPAVLAALQHS 265
DB 197 GSIASGWAQAQPSHAPHELWISGAALGGVVPNTAAVATKVSQGVFAGLVPAGITGFQNG 256
QY 266 YPLGARVIRKEANDSGRQLLEQLTMTTVDVIRMAGDMGDFLDEPLED-----ILS 318
DB 257 YDEVSEIISAHLPVSKVEBNKANSQCIVADALGVAFQDI-----FEDFTTGTGILE 309
QY 319 TPEISHVFGDTKLGSAVTPPVLIVQAVHYDLIDVSDIDLADSY-TAGGANVTYHRDLF 377
DB 310 VPEVAARLNQNLGSLVPEFPYIYHVGDEVAPVEKADTLVSSNCNAGIKSLEYHKDIL 369
QY 378 SEHVSLLPLSDPMTLRL 395
DB 370 DEHLSEEVFGSAGATVWL 387
```

```
RESULT 8
ADP98845
ID ADP98845 standard; protein; 460 AA.
XX AC ADP98845;
XX DT 23-SEP-2004 (first entry)
XX DE C. albicans specific gene, orf6.2643, protein sequence.
XX KW Diploid fungal cell; allele; gene disruption cassette;
XX KW promoter replacement fragment; antifungal; fungicide; gene therapy;
XX KW infection; Candida albicans.
XX OS Candida albicans.
XX PN WO2004056965-A2.
XX PD 08-JUL-2004.
XX PF 19-DEC-2003; 2003WO-US040618.
XX PR 19-DEC-2002; 2002US-0434832P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PA (ELIT-) ELITRA PHARM LTD.
XX PI Roemer T, Jiang B, Boone C, Bussey H;
XX WPI; 2004-500296/47.
XX N-PSDB; ADP98535.
XX PT Constructing a strain of diploid fungal cells in which both alleles of a
XX PT gene are modified comprises modifying the alleles of a gene in the fungal
XX PT cells by recombination using a gene disruption cassette and a promoter
XX PT replacement fragment.
XX PS Claim 44; SEQ ID NO 7020; 163pp; English.
XX CC The invention relates to a novel method for constructing a strain of
XX CC diploid fungal cells in which both alleles of a gene are modified. The
XX CC method comprises modifying the alleles of a gene in diploid fungal cells
XX CC by recombination using a gene disruption cassette and a promoter
XX CC replacement fragment. The invention further comprises: assembling a
XX CC collection of diploid fungal cells each of which comprises modified
XX CC alleles of a different gene; a strain of diploid fungal cells comprising
XX CC modified alleles of a gene, where the first allele of the gene is
XX CC inactivated by a gene disruption cassette comprising a nucleotide
XX CC sequence encoding an expressible selectable marker; and the expression of
XX CC the second allele of the gene is regulated by a heterologous promoter
XX CC that is operably linked to the coding region of the second allele of the
XX CC gene, and where the gene encodes the polypeptide mentioned above; a
XX CC collection of diploid fungal strains comprising the diploid strains cited
XX CC above, where substantially all the different genes that encode the above
XX CC amino acid sequences are modified and are present in different diploid
XX CC strains in the collection; a nucleic acid molecule microarray comprising
XX CC nucleic acid molecules, where each nucleic acid molecule comprises a
XX CC nucleotide sequence that is hybridizable to a target nucleotide sequence
XX CC comprising any of the 110 nucleotide sequences listed in the
XX CC specification (ADP98516-ADP98825); identifying a gene that is essential
XX CC to the survival or growth of a fungus, that contributes to the virulence
XX CC and/or pathogenicity of a fungus, or that contributes to the resistance
XX CC of a diploid fungus to an antifungal agent; identifying an antifungal
XX CC agent that inhibits the growth of a diploid fungus, or a therapeutic
XX CC agent for treatment of a mammalian disease; correlating changes in the
XX CC levels of proteins or gene transcripts with the inhibition of growth or
XX CC proliferation of a diploid fungal cell; a purified or isolated nucleic
XX CC acid molecule comprising a nucleotide sequence encoding a gene product
XX CC required for proliferation of Candida albicans, where the gene product
XX CC consists of any of the above-mentioned amino acid sequences; a vector
XX CC comprising a promoter operably linked to the nucleic acid molecule cited
XX CC above; a host cell containing the vector; a purified or isolated
```

CC polypeptide comprising any of the 61 amino acid sequences given in the
CC specification (ADP96718-ADP96778); a fusion protein comprising a fragment
CC of a first polypeptide fused to a second polypeptide, the fragment
CC consisting of at least 6 consecutive residues of any of ADP98826-ADP99135
CC ; producing a polypeptide; identifying a compound which modulates the
CC activity of a gene product encoded by a nucleic acid comprising any of
CC ADP98516-ADP98825; eliciting an immune response in an animal; a strain of
CC Candida albicans, where a first allele of a gene comprising any of
CC ADP98516-ADP98825 is inactive and a second allele of the gene is under
CC the control of a heterologous promoter; identifying a compound or binding
CC partner that binds to the polypeptide comprising any of ADP98826-
CC ADP99135, or its fragment; identifying a compound having the ability to
CC inhibit growth or proliferation of Candida albicans; inhibiting growth or
CC proliferation of Candida albicans cells; manufacturing an antimycotic
CC compound; treating an infection of a subject by Candida albicans;
CC preventing or containing contamination of an object by Candida albicans,
CC or for preventing or inhibiting formation on a surface of a biofilm
CC comprising Candida albicans; a pharmaceutical composition comprising a
CC therapeutic amount of an agent which reduces the activity or level of a
CC gene product encoded by a nucleic acid comprising any of ADP98516-
CC ADP98825 in a pharmaceutical carrier; an antibody preparation which binds
CC the polypeptide; methods for evaluating a compound against a target gene
CC product encoded by any of ADP98516-ADP98825; identifying an antimycotic
CC compound; a computer or a computer readable medium that comprises at
CC least one of the nucleotide sequences mentioned in the specification or
CC at least one amino acid sequence selected from ADP98826-ADP99135; a
CC method assisted by a computer for identifying a putatively essential gene
CC of a fungus; and a protein array comprising proteins, where at least one
CC protein comprises an amino acid sequence or a portion of an amino acid
CC sequence selected from ADP98516-ADP98825. The novel methods and
CC compositions have fungicide activity. The compositions may be used in
CC gene therapy. The composition and methods are useful for drug screening
CC purposes or for diagnosing, preventing or treating infections associated
CC with Candida albicans. These may also be used for constructing strains
CC useful for identification and validation of gene products as effective
CC targets for therapeutic intervention, for identifying and validating gene
CC products as effective targets for therapeutic intervention, and for
CC collecting identified essential genes. This sequence represents the
CC protein of a Candida albicans fungal specific gene of the invention.
CC NOTE: This sequence was downloaded from an electronic sequence listing
CC provided on the WIPO website.
XX
SQ Sequence 460 AA;

Query Match 21.1%; Score 491.5; DB 8; Length 460;
Best Local Similarity 30.5%; Pred. No. 1.9e-35;
Matches 120; Conservative 78; Mismatches 162; Indels 33; Gaps 10;
QY 30 PLLPSDDPFYFPAGYQHAVPGTVLSRSDVELAFMG-LIPQVY-TATQLLYRTTMYGNP 87
DB 19 PTKPSSDPFYPNPKGFEKAAVGDILQSRTPKSIITGRFAPLKIQNSWQLLVRSSEDSFGNP 78
QY 88 EATVTTVIIPAELAPGOTCPLLSYQCAIDAMSSRCPPSYALRRRAKALGSLTQMELMIS 147
DB 79 NAIIVTVIEPVNADPK---IASYQVFEDAAXADCAPSALQPGSLDTFTVQAENYLMWA 135
QY 148 AALAEQWAVSPDHEPGKGLWGSYPBPGYRVLDGIRAAALNSRV-GLSPATPIGLMGYSG 206
DB 136 PLLDQGVYVSPDYEGRKSTFTTGKSGQAVLNSIRATLKSSKITNIKEDAKVWVMGYSG 195
QY 207 GGLASWAAEACEYAPDL--DIVGAVLSPVGDGLHTFRRLNGTLGALPALVVAALQH 264
DB 196 GSLASGAAALQPSYAPELSSLLGAALGGFVTNITATAQAADGTVFAGIVANALGGVAN 255
QY 265 SYPLGARVKEHANDREGROLLEQLTEMTTVDVIRWAGRMDGDFLDEPL----- 313
DB 256 EYEPFKSIQ--SDTKKSVDFDSDHCLADGVI-----DYINTSFLTGDKNIFKTKG 305
QY 314 EDILSTPEISHVFGDTKL---GSAVTPPVIQVAHVHYLDVSDIDALADSVTAGG-AN 369
DB 306 WDLKSPSTIAKIVEDNGLYQKQVLPKIPFVYVHGSIDQIVPVIYVKKTYQNWCEGGISS 365
QY 370 VTYHRDLFSEHVLHPLSAPMTLRWLTDRPAGK 402

Db 366 LEFAEDGTNGHLTETVVGAPAAALTWIIDRFNGK 398
RESULT 9
ID ABP73893
XX ABP73893 standard; protein; 468 AA.
AC ABP73893;
XX
DT 30-JAN-2003 (first entry)
XX
DE Candida albicans essential protein SEQ ID NO 7730.
XX
KW Fungus; yeast; tetracyclin; promoter; GRACE strain; biosynthesis;
KW signal transduction; DNA replication; cell division; growth;
KW proliferation; Candida albicans; fungicide; antifungal.
XX
OS Candida albicans.
XX
PN WO200253728-A2.
XX
PD 11-JUL-2002.
XX
PF 26-DEC-2001; 2001WO-US049486.
XX
PR 29-DEC-2000; 2000US-0259128P.
PR 20-FEB-2001; 2001US-00792024.
PR 22-AUG-2001; 2001US-0314050P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;
XX
DR WPI: 2002-566694/60.
DR N-PSDB; AB232443.
XX
PT Constructing strains for identifying gene products as effective targets
PT for therapeutic intervention, by inactivating in the strain one allele of
PT a gene and placing other allele of the gene under conditional expression.
XX
PS Claim 44; SEQ ID NO 7730; 167pp + Sequence Listing; English.
XX
CC The invention relates to constructing (M1) a strain of diploid fungal
CC cells in which both alleles of a gene are modified, comprising modifying
CC one allele by insertion or replacement by a cassette having an
CC expressible selectable marker and modifying other allele by
CC recombination, of a promoter replacement fragment with a heterologous
CC promoter, so that expression of the second allele is regulated by the
CC promoter. (M1) is useful for constructing a strain of diploid fungal
CC cells in which both alleles of a gene are modified. The diploid fungal
CC cells having both alleles modified are useful for identifying a gene that
CC is essential to the survival or growth of a fungus, a gene that
CC contributes to the virulence and/or pathogenicity of a fungus, a gene
CC that contributes to the resistance of a diploid fungus to an antifungal
CC agent, an antifungal agent that inhibits the growth of a diploid fungus
CC and for identifying a therapeutic agent for treatment of a mammalian
CC disease. (M1) is useful for identifying a compound which modulates the
CC activity of a gene product, preferably enzymatic activity, carbon
CC compound catabolism, biosynthetic, transporter, transcriptional,
CC translational, signal transduction, DNA replication and cell division
CC activity. The method is useful for identifying a compound having the
CC ability to inhibit growth or proliferation of C. albicans cells and for
CC treating infection by C. albicans. The present sequence is that of an
CC essential Candida albicans protein used in the method of the invention.
CC Note: The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office
XX
SQ Sequence 468 AA;
Query Match 18.8%; Score 439; DB 5; Length 468;
Best Local Similarity 29.7%; Pred. No. 1.1e-30;

DR WPI: 2003-259157/26.
DR N-PSDB; ABX94752.
XX
PT New polypeptide with lipase/acetyltransferase activity isolated from
PT Candida parapsilosis is useful as catalyst in acetyl transferase
PT reactions in synthesis of esters, thioesters and lactones.
XX
PS Claim 3; Page 16-17; 28pp; German.
XX
CC This invention describes a novel lipase/acyltransferase isolated from
CC Candida parapsilosis. The polypeptides of the invention are used as
CC catalysts in acyl transfer reactions, particularly alcoholysis of an
CC ester, thioester, thiolysis of a ester, aminolysis of esters with hydroxy
CC amines of hydrazines, reactions of esters with hydrogen peroxide or
CC synthesis of esters, thioesters or lactones through alcoholysis. This
CC sequence represents a lipase/acyltransferase described in the invention
XX
SQ Sequence 465 AA;
Query Match 18.8%; Score 437.5; DB 6; Length 465;
Best Local Similarity 30.0%; Pred. No. 1.5e-30;
Matches 127; Conservative 70; Mismatches 159; Indels 67; Gaps 12;
QY 30 PLLPSDDPFYPPAGYQHVPCTVLRSDVELAFMGLIPQPVTT-----ATQLLYR 79
Db 21 PKPSQDDFTYPPQGYEAQPLGSLIKTRNV-----PNPLTNVFTPVKVQNAWQLLVR 72
QY 80 TTNMYGNPEATVTVIVPAELAPGQTCPLLSYQCAIDAMSSRCFSPSALRRRAKALGSIT 139
Db 73 SEDTGNPNAIVTTIIQPFNAKKDK---LVSYQTPEDSKGLDCAPSALQYQSDISTLT 129
QY 140 QMELLMISAALAEAGWAVSPDHEGPKGLWGSPYEPGYRVLGIRAAALNSERV-GLSPATP 198
Db 130 QGEMYIISALLDQGYVVTVPDYEGRPKSTFTVGLQSGRATLNSLRATLKSGLTGVSSDAE 189
QY 199 IGLWYSGGGLASAAAEACGEYAPDL--DIVGAVLGSPVGLDHTFRRLNGLTLAGLPA 256
Db 190 TLLWYSGGSLASGAAAIQKEYAPELSKNLLGAALGGFVTNITATAEAVDSGPFAGIIS 249
QY 257 LVVAALQHSYPGLARVIKEHANDEGRQLLEQLEMTTV-----DAVIRMAGR-- 303
Db 354 VNSRKTFOQCDWGLKSGEYNEDLTNGHITESIVGAPAAALTWIIINRFGQPPVDCQHN 413
QY 410 RTT 412
Db 414 RAS 416
RESULT 12
ID ADP08158
XX ADP08158 standard; protein; 465 AA.
XX AC ADP08158;
XX DT 29-JUL-2004 (first entry)
DE Lipase/acyltransferase SEQ ID 2.
XX Lipase; acyltransferase; enzyme.
XX Candida parapsilosis.
XX CA2403025-A1.
XX 08-APR-2004.

XX 15-OCT-2002; 2002CA-02403025.
XX 08-OCT-2002; 2002US-0416987P.
XX (COGN-) COGNIS DEUT GMBH & CO KG.
XX Dubreucq E, Moulin G, Bigey F, Weiss A;
XX WPI; 2004-390791/37.
DR N-PSDB; ADP08157.
XX Novel isolated polypeptide with lipase/acyltransferase activity useful as
PT catalysts in chemical and biochemical reactions.
XX Claim 1; SEQ ID NO 2; 46pp; English.
XX The present invention relates to novel polypeptides (I) with
CC lipase/acyltransferase activity from Candida parapsilosis (ADP08158 and
CC ADP08160). (I) are useful as catalysts in acyltransfer reactions,
CC especially the reactions which are chosen from alcoholysis of esters,
CC alcoholysis of thioesters, thiolysis of esters, aminolysis of an ester
CC with hydroxylamines or hydrazines, reaction of an ester with hydrogen
CC peroxides and enantioselective synthesis of esters, thioesters and
CC lactones by alcoholysis. The present sequence is the coding sequence for
CC a lipase/acyltransferase.
XX
SQ Sequence 465 AA;
Query Match 18.8%; Score 437.5; DB 8; Length 465;
Best Local Similarity 30.0%; Pred. No. 1.5e-30;
Matches 127; Conservative 70; Mismatches 159; Indels 67; Gaps 12;
QY 30 PLLPSDDPFYPPAGYQHVPCTVLRSDVELAFMGLIPQPVTT-----ATQLLYR 79
Db 21 PKPSQDDFTYPPQGYEAQPLGSLIKTRNV-----PNPLTNVFTPVKVQNAWQLLVR 72
QY 80 TTNMYGNPEATVTVIVPAELAPGQTCPLLSYQCAIDAMSSRCFSPSALRRRAKALGSIT 139
Db 73 SEDTGNPNAIVTTIIQPFNAKKDK---LVSYQTPEDSKGLDCAPSALQYQSDISTLT 129
QY 140 QMELLMISAALAEAGWAVSPDHEGPKGLWGSPYEPGYRVLGIRAAALNSERV-GLSPATP 198
Db 130 QGEMYIISALLDQGYVVTVPDYEGRPKSTFTVGLQSGRATLNSLRATLKSGLTGVSSDAE 189
QY 199 IGLWYSGGGLASAAAEACGEYAPDL--DIVGAVLGSPVGLDHTFRRLNGLTLAGLPA 256
Db 190 TLLWYSGGSLASGAAAIQKEYAPELSKNLLGAALGGFVTNITATAEAVDSGPFAGIIS 249
QY 257 LVVAALQHSYPGLARVIKEHANDEGRQLLEQLEMTTV-----DAVIRMAGR-- 303
Db 250 NALAGIGNEYP-----DFKNYLLKKVSPLLSITYRLGNTHCLLDGGIAYFGKSF 298
QY 304 -----DMGDFLD-EPLEDILSTPEISHVFGDTKLGSVPTPPVLIVQAVHDYLDV 353
Db 299 FSRIRYFPDGDWLVNQEPKTILODNGLVYQPKD-----LTPQIPLFIYHGTLDIAVPI 353
QY 354 SDIDALADSYTAGG-ANVTYHRDLFSEHVSHPLSAPMTLRLWLTDRFAGKPLTD---HRV 409
Db 354 VNSRKTFOQCDWGLKSGEYNEDLTNGHITESIVGAPAAALTWIIINRFGQPPVDCQHN 413
QY 410 RTT 412
Db 414 RAS 416
RESULT 13
ID ABG74897
XX ABG74897 standard; protein; 471 AA.
XX AC ABG74897;
XX DT 03-JUL-2003 (first entry)

XX C. parapsilosis lipase/acyltransferase SEQ ID 4.
DE Lipase; acyltransferase; catalyst; acyl transfer; alcoholysis; ester;
KW thioester; thiolysis; aminolysis; hydrazine; lactone; enzyme.
XX Candida parapsilosis.
XX EPI275711-A1.
PN 15-JAN-2003.
XX 11-JUL-2001; 2001EP-00401855.
PF 11-JUL-2001; 2001EP-00401855.
PR 11-JUL-2001; 2001EP-00401855.
XX (COGN-) COGNIS DEUT GMBH & CO KG.
PA Dubreucq E, Bigey F, Moulin G, Weiss A;
PI WPI; 2003-259157/26.
DR N-PSDB; ABX94753.
XX New polypeptide with lipase/acyltransferase activity isolated from
PT Candida parapsilosis is useful as catalyst in acetyl transferase
PT reactions in synthesis of esters, thioesters and lactones.
XX Claim 21; Page 19-20; 28pp; German.
PS This invention describes a novel lipase/acyltransferase isolated from
CC Candida parapsilosis. The polypeptides of the invention are used as
CC catalysts in acyl transfer reactions, particularly alcoholysis of an
CC ester, thioester, thiolysis of an ester, aminolysis of esters with hydroxy
CC amines of hydrazines, reactions of esters with hydrogen peroxide or
CC synthesis of esters, thioesters or lactones through alcoholysis. This
CC sequence represents a lipase/acyltransferase described in the invention
XX Sequence 471 AA;
SQ

Query Match 18.8%; Score 437.5; DB 6; Length 471;
Best Local Similarity 30.0%; Pred. NO. 1.6e-30;
Matches 127; Conservative 70; Mismatches 159; Indels 67; Gaps 12;

QY 30 PLLSDDPFPYPAGYQHAVGTGLRSRDVELAFMGLIPOPVT-----ATQLLYR 79
DB 21 PKPSQDDFYTPPGYEAQPLGSLIKTRNV-----PNLTNVFTFPVKQNAWQLLVR 72
QY 80 TTNMYGNPEATVTTVIVPAELAPGQTCPLLSYQCAIDAMSSRCPPSYALRRRAKALGSLT 139
DB 73 SEDTFGNPNAIVTTIIQFPNAKKDK--LVSYQTFEDSGKLDCAPSVAIQGSDISTLTT 129
QY 140 QMELLMISAALAEAWAVSPDHEGPKLWGSFPGYRVLDGIRAAALNSERV-GLSPATP 198
DB 130 QGEMYYISALLDQGYVVVTPDYGPKSTFTVGLQSGRATLSRLATLKSGLTGVSSDAE 189
QY 109 IGLMGYSGGGLASAAAEACGEYAPDL--DIVGAVLGSVPDGLGHTFRLNGTLLAGLPA 256
DB 190 TLLMGYSGGSLASGAAAIQKEYAPELSKNLLGALGGFVTNITATAEAVDSGPFAGIIS 249
QY 257 LVVAALQHSYFGLARVIKEHANGDEGRQLLEQTEMVTV-----DAVIRMAGR-- 303
DB 250 NALAGIGNEYP-----DFKNYLLKKVSPLLSITYRLGNTHCLLDGGIAVFGKSF 298
QY 304 -----DMGDFLD-EPLSDILSTPEISHVFGDTKLGSVAPPPVULIVCAVHDYLDV 353
DB 299 FSRIRYFPDGDVLNQEPIKTIQDNLGLVYQPKD-----LTPQIPLFIYHGTLDIAVPI 353
QY 354 SDIDALADSYTAGG-ANVTYHRDLFSEHVSLLPSAPMTLRLWLTDRFAGKPLTD---HRV 409
DB 354 VNSRKTFOQCDWGLKSEYNEDLTNGHITESIVGAPAAALTWIINRNGQPPVDGCCQHV 413
QY 410 RTT 112

Db 414 RAS 416

RESULT 14
ADP08160
ID ADP08160 standard; protein; 471 AA.
XX ADF08160;
XX 29-JUL-2004 (first entry)
XX Lipase/acyltransferase SEQ ID 4.
XX Lipase; acyltransferase; enzyme.
XX Candida parapsilosis.
OS Synthetic.
XX CA2403025-A1.
XX 08-APR-2004.
XX 15-OCT-2002; 2002CA-02403025.
XX 08-OCT-2002; 2002US-0416987P.
XX (COGN-) COGNIS DEUT GMBH & CO KG.
XX Dubreucq E, Moulin G, Bigey F, Weiss A;
XX WPI; 2004-390791/37.
XX N-PSDB; ADP08159.
XX Novel isolated polypeptide with lipase/acyltransferase activity useful as
PT catalysts in chemical and biochemical reactions.
XX Claim 8; SEQ ID NO 4; 46pp; English.
XX The present invention relates to novel polypeptides (I) with
CC lipase/acyltransferase activity from Candida parapsilosis (ADP08159 and
CC ADP08160). (I) are useful as catalysts in acyltransfer reactions,
CC especially the reactions which are chosen from alcoholysis of esters,
CC alcoholysis of thioesters, thiolysis of esters, aminolysis of an ester
CC with hydroxylamines or hydrazines, reaction of an ester with hydrogen
CC peroxides and enantioselective synthesis of esters, thioesters and
CC lactones by alcoholysis. The present sequence is a modified
CC lipase/acyltransferase which comprises a 6-His-peptide at the C-terminal
CC end.
XX Sequence 471 AA;
SQ

Query Match 18.8%; Score 437.5; DB 8; Length 471;
Best Local Similarity 30.0%; Pred. NO. 1.6e-30;
Matches 127; Conservative 70; Mismatches 159; Indels 67; Gaps 12;

QY 30 PLLSDDPFPYPAGYQHAVGTGLRSRDVELAFMGLIPOPVT-----ATQLLYR 79
DB 21 PKPSQDDFYTPPGYEAQPLGSLIKTRNV-----PNLTNVFTFPVKQNAWQLLVR 72
QY 80 TTNMYGNPEATVTTVIVPAELAPGQTCPLLSYQCAIDAMSSRCPPSYALRRRAKALGSLT 139
DB 73 SEDTFGNPNAIVTTIIQFPNAKKDK--LVSYQTFEDSGKLDCAPSVAIQGSDISTLTT 129
QY 140 QMELLMISAALAEAWAVSPDHEGPKLWGSFPGYRVLDGIRAAALNSERV-GLSPATP 198
DB 130 QGEMYYISALLDQGYVVVTPDYGPKSTFTVGLQSGRATLSRLATLKSGLTGVSSDAE 189
QY 199 IGLMGYSGGGLASAAAEACGEYAPDL--DIVGAVLGSVPDGLGHTFRLNGTLLAGLPA 256
DB 190 TLLMGYSGGSLASGAAAIQKEYAPELSKNLLGALGGFVTNITATAEAVDSGPFAGIIS 249
QY 257 LVVAALQHSYFGLARVIKEHANGDEGRQLLEQTEMVTV-----DAVIRMAGR-- 303

THIS PAGE BLANK (U8PT0)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 01:42:44 ; Search time 36.6186 Seconds
(without alignments)
909.197 Million cell updates/sec

Title: US-10-617-038-8
Perfect score: 2332
Sequence: 1 MVEPGLAGATGAEWIGRPP.....RLAVIAKVITGRKLSRRPL 446

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A-COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B-COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	504.5	21.6	422	4	US-09-248-796A-16480
2	482.5	20.7	424	4	US-09-248-796A-16475
3	467.5	20.0	461	4	US-09-248-796A-16479
4	437.5	18.8	467	4	US-09-248-796A-16476
5	393	16.9	437	4	US-09-248-796A-16483
6	383	16.4	462	1	US-08-458-023B-2
7	383	16.4	463	3	US-09-111-556A-2
8	383	16.4	463	3	US-08-360-758-2
9	363	15.6	266	4	US-09-248-796A-16484
10	351	15.1	285	4	US-09-248-796A-16474
11	335.5	14.4	320	3	US-09-036-987A-14
12	335.5	14.4	320	3	US-09-370-700-14
13	335.5	14.4	320	4	US-09-603-207-14
14	274.5	11.8	435	4	US-09-602-787A-238
15	274.5	11.8	435	4	US-09-602-777A-254
16	266	11.4	289	4	US-09-248-796A-16473
17	261.5	11.2	228	4	US-09-248-796A-16477
18	219.5	9.4	203	4	US-09-248-796A-16486
19	205.5	8.8	293	4	US-09-248-796A-16482
20	181.5	7.8	402	4	US-09-328-352-4980
21	149	6.4	265	4	US-09-248-796A-16478
22	124	5.3	424	4	US-09-328-352-4377
23	112	4.8	3724	2	US-08-804-227C-10
24	112	4.8	3724	2	US-08-804-198-4
25	109.5	4.7	660	4	US-09-902-540-14368
26	107	4.6	489	4	US-09-328-352-4634

28	107	4.6	1436	2	US-08-652-971-2	Sequence 2, Appli
29	107	4.6	1436	2	US-08-991-258A-2	Sequence 2, Appli
30	107	4.6	1436	2	US-08-769-399-2	Sequence 2, Appli
31	107	4.6	1436	3	US-08-991-953A-2	Sequence 2, Appli
32	106.5	4.6	194	4	US-09-248-796A-16485	Sequence 16485, A
33	106	4.5	711	4	US-09-252-991A-22865	Sequence 22865, A
34	105.5	4.5	600	4	US-09-287-849-22	Sequence 22, Appli
35	105.5	4.5	2284	4	US-09-252-991A-23547	Sequence 23547, A
36	105	4.5	884	4	US-09-252-991A-26707	Sequence 26707, A
37	105	4.5	3519	3	US-09-428-517-4	Sequence 4, Appli
38	104.5	4.5	470	4	US-09-266-965-118	Sequence 118, App
39	104	4.5	761	1	US-07-906-395-2	Sequence 2, Appli
40	104	4.5	761	1	US-08-192-632-2	Sequence 2, Appli
41	104	4.5	761	1	US-08-710-676-2	Sequence 2, Appli
42	104	4.5	761	3	US-09-099-902B-2	Sequence 2, Appli
43	104	4.5	761	5	PCT-US93-06080-2	Sequence 2, Appli
44	104	4.5	1047	1	US-08-190-687B-8	Sequence 8, Appli
45	103.5	4.4	4472	2	US-08-804-227C-2	Sequence 2, Appli

RESULT 1
US-09-248-796A-16480
; Sequence 16480, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16480
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16480

Query Match 21.6%; Score 504.5; DB 4; Length 422;
Best Local Similarity 32.1%; Pred. No. 3.1e-39;
Matches 127; Conservative 75; Mismatches 159; Indels 35; Gaps 10;
Qy 28 VRLLPSDDPYFPAGYQHAVPGTVLRSRDVELAFMGLIPQVTA*-----QLL 77
Db 30 ILPTKPSNDPFPYADAPGKAAVGDILQSRKT-----PKPITGVFVVKIQNSWQLL 81
Qy 78 YRTNMYGNPATVTVIVPAELAPGQCPLLSYQCAIDAMSSRCFSPVALRRRAKALGS 137
Db 82 VRSEDFGNPNVITVTMPEPNADPSK---LASQVFEDSAKADCAPSALFQGDVTTI 138
Qy 138 LTQEMLLMISAALAEQMAVSPDHEGPKGLWGSPEYGPVRLDGIKRAALNSRV-GLSPA 196
Db 139 ATQVETYLAPLDGQYVVSFDEGPKLFTTVGKQSGQAVLNSRAALKSKGKINTLAED 198
Qy 197 TPGLWGYSGGLASAWAAEACGEYAPDL--DIVGAVLGSVPVGLDGLFRRRLNGTLTLAGL 254
Db 199 AKVVMWGYSGGLASGMAAALQPNYAPELGGLNLLGAALGGFVNTATATAEATDGVFAGI 258
Qy 255 PALVVAALQHSYPLGARVIRKEHANEGRQ-LLEQLTETMTTVDATVRMAGRDGDFLDEPL 313
Db 259 MANALGGVANEYEPFKQILQ---NDTDKQSFQDNNHCLADGVNITCKHFLSGTNKIF 315
Qy 314 E---DILSTPETSISHVFGDTKL---GSAVTPPVLIVQAVHDYLDIVSDIDALADSY-TAG 366
Db 316 KSGWNILKNPTISKIVEDNGLVYQKQLPKPILLYHGAIDQIVFVNVKTYQNKWDAG 375

ALIGNMENTS

Db 77 SEDSGFNPAFTTTIQKNADPSK---VVSQNWEDASINCSYSGQLGAPLSTILT 133
Qy 140 QMELLMISAALAEAGWAVSPDHEGKGLMGSPYEGYRVLGIRAAALNSERV-GLSPATP 198
Db 134 QLDWTFIVPLKSGYVVLDPYEGPKSTFGVGRQSKATLDSIKAVLTKTKDFSGINDDAK 193
Qy 199 IGLWYSGGGLASAWAAEACGEYAPDL--DIVGAVLGSVGLDGLHTFRLNGTLTLAGLPA 256
Db 194 VVLWYSGGSGFASGAAVLQPEYAPBLKDNLLICAAALGFAANLTGIAESVDGEVFGTIP 253
Qy 257 LVVAALQHSYPGLARVIRKEHANDEGRQLLEOLTMTTVDVAVR----- 299
Db 254 LALNGIANEYDPDKRLYEVEVPGAKADLQKAENCLASLISYPMYQFTGPRRVFEKG 313
Qy 300 ---MAGRDMDGLDEPLEDILSTPEISHVFGDTKLGSAVPTPVLIVQAVHDYLDVSDI 356
Db 314 WSLLEDKTKTGLDNLTLALSKEH-----MPQIPFVYHGTIDKIPIKDS 360
Qy 357 DALADSYTAGG-ANVTYHRDLFSEHVSHPGLSAPMTLRLWLTDRFAGKPLTDHRVTTWPT 415
Db 361 IKIYNKWCWIGSGFSEFSDKLNGHTTETVVGAPAAALTWIDARFAGKPAVEGCLFTTRAS 420
Qy 416 IENPWTYACMARLAVIAAKVI 436
Db 421 NF---LYPNISSSAALYFKGI 438

RESULT 5
US-09-248-796A-16483
; Sequence 16483, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248, 796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16483
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16483

Query Match 16.9%; Score 393; DB 4; Length 437;
Best Local Similarity 28.8%; Pred. No. 1.2e-28;
Matches 115; Conservative 74; Mismatches 171; Indels 40; Gaps 10;

Qy 44 GYQHAVPGTVLRSRDELAFMGL-IPQPV-TATOLLRYTNNYGNPEATVTVVPAELA 101
Db 1 GYESAKLGILKLRKTPGKISSLFIPEVKNSQWLLVRSSEDSFGNAAAIVTTVIEPFNAD 60
Qy 102 PGQTCPLLSYQCAIDAMSSRCFSPYALRRRAKALGSLTQWELMISAALAEAGWAVSPDH 161
Db 61 PSK---VVSQSWEDANIECSYSGWQFGAPLSSVQTQVDHFIIVPLDKGCFVVLDPY 117
Qy 162 EGPGLWGSYPYEGYRVLGIRAAALNSERV-GLSPATPIGLMGYSGGGLASAWAABACGE 220
Db 118 EGPKSTFGVGRQSGKATLDSIKAVLTKTKDFSGINDDAQVAMWYSGGTIAAGWAATLQPK 177
Qy 221 YAPDL--DIVGAVLGSVGLDGLHTFRLNGTLTLAGLPAALVVAALQHSYPGLARVIRKEHAN 278
Db 178 YAOELKKNLTGAALGGFVINITATABATDGTFLPAGLIPNALNGLANEFPDFKRMVYEVVE 237
Qy 279 DEGRQLLEOLTMTTVDVIRMAGRDM--GD-----FLDEPLEDILSTPEISHVFG 327
Db 238 KRYEGALQOQGTQCLGAILHFAFDQVFTGDHRYFEQGYGLLEE-----EVFN 285

Qy 328 DTKLGS-----VPTPPVLIVQAVHDYLDVSDIDALADSYTAGGA-NVTYHRDLFS 378
Db 286 RTILGNSLLYMDQEYLPDIPFVYHGSLEDGIVPIPDVHGVYKNWCDWGDSEFAEDSLN 345
Qy 379 EHVSLHPLSAPMTLRLWLTDRFAGKPLTDHRVTTWPTIFN 418
Db 346 GHLTEIVVGAPAAITWLDARFDGQPVVEGCKKTRITNFS 385

RESULT 6
US-08-458-023B-2
; Sequence 2, Application US/08458023B
; Patent No. 5667990
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Yoder, Wendy
; APPLICANT: Takagi, Shinobu
; APPLICANT: Boominathan, Karuppan C.
; TITLE OF INVENTION: ASPERGILLUS EXPRESSION SYSTEM
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5667990o No. 5667990disk of No. 5667990th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,023B
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4086.010-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 462 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-458-023B-2

Query Match 16.4%; Score 383; DB 1; Length 462;
Best Local Similarity 27.2%; Pred. No. 1.2e-27;
Matches 127; Conservative 64; Mismatches 210; Indels 66; Gaps 10;

Qy 7 LAGATCAEWIGRPHEELQKVRPLPSDDPFYPPAGYQHAVPGTVLRSRDELAFMGL 66
Db 12 LAAATAAV-LAAPAAETLDRRAALPNPYDDPFYTPPSNIGTFAGKQVIOQRKV----- 63
Qy 67 IQQPV-----TATOLLRYTNNYGNPEATVTVVPAELAPCQTCPLLSYQCAIDAMSS 120
Db 64 -PTDIGNANNAASFOLQYRTTNTQNEAVADVATVWIPAK--PASPKIFSYQVYEDATAL 120
Qy 121 RCFPSYALRRRAKALGSLTQM--ELLMISAALAEAGWAVSPDHGEGPKGLMGSPYEGYR 178
Db 121 DCAPSYLYLTGLDQPNKVTAVLDTPIIIGWALQOQYVYVSSDHGEGKAAFIAGYEGMAI 180
Qy 179 LDGIRAAALNSERVGLSPATPIGLMGYSGGGLASAWAAEACGEYAPDLDIVGAVLGSVGD 238
Db 181 LDGIRALKNYQ--NLPSDSKVALEGYSGGAHATVWATSLAESYAPELNIVGASHGTPVS 238
Qy 239 LGHTFRLNGTLTLAGLPAALVVAALQHSYPGLARVIRKEHANDEGRQLLEOLTMTTVDVAVI 298

```
Db 239 AKDTFTLNGGPFAGFALAGVSLAHPDMESFIEARLNNAKQRTLKQI----- 288
Qy 299 RMAGDMGDFLDE-----PLEDILSTPEISHVFGDTKLGS-----AVPT 337
Db 289 ----RGRGFCPLPQVVLTYPFNLVFLVNDTNLLNEAPIASILKQETVVQABASVTVSPK 344
Qy 338 PVLIVQAVHDYLDVSDIDALADSYTAGGANVTYHRDLFSEHVSLLHPLSAPMTLRWLT 397
Db 345 PPRFIWHAIPDEIVPYQPAATYVKEQCAKANINPSPYIAEHLTAETIFGLVPSLWFIK 404
Qy 398 RFAGKPLTDHRVRTTWTPTIFNPMTYAGMARLAVIAAKVITGRKLSRR 444
Db 405 AFDG-----TTPKVICGTPIPALAGITTPSADQVGLGDLANQ 441

RESULT 7
US-09-111-556A-2
; Sequence 2, Application US/09111556A
; Patent No. 6020180
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Pathar, Shamkant A
; APPLICANT: Egel-Mitani, Michi
; APPLICANT: Borch, Kim
; APPLICANT: Clausen, Ib G
; APPLICANT: Hansen, Mogens T
; TITLE OF INVENTION: C. ANTARCTICA LIPASE AND LIPASE VARIANTS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESS: No. 6020180o No. 6020180disk of No. 6020180th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,556A
; FILING DATE: 22-DEC-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK PCT/DK93/00225
; FILING DATE: 03-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3748.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 463 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-111-556A-2

Query Match 16.4%; Score 383; DB 3; Length 463;
Best Local Similarity 27.2%; Pred. NO. 1.2e-27;
Matches 127; Conservative 64; Mismatches 210; Indels 66; Gaps 10;

Qy 7 LAGATGAEMIGRPPEELQKVRPLLPSSDDPPFPYFPAGYQHAVPGTVLRSDELAFNGL 66
Db 12 LAATAAV-LAAPAETLDRALPNPYDDPPFTPSNIGTFAGQVQSRKV----- 63
Qy 67 IPQPV-----TATOLLYRTTNMYGNPEATVTIVPAELAPGQTCPLLSYQCAIDAMSS 120
```

```
Db 64 -PTDIGNANNAASQOLQYRTINTONEBAVDVATVWIPAK--PASPPKIFSQVYEDATAL 120
Qy 121 RCFPSYALRRRAKALGSLTQM--ELLMISAALAGWAVSPDHGPKGLWGSPEYEGYRV 178
Db 121 DCAPSYLVLTGLDQPNKVTAVLDTPIIGMALQQGYVWSSDHGEGFAAFIAGYEEMAI 180
Qy 179 LDGIRALNSERVCLSPATPIGLWGYSGGGLASAWAAEACGEYAPDLDIVCAVIGSPVGD 238
Db 181 LDGIRALKNYQ--NLPSDSKVALEGYSGGAHATVWATSLAESYAPELNIVGASHGGTPTS 238
Qy 239 LGHTFRRLNGTLLAGLPALVVAALQHSYPGLARVIKEHANDEGRQLLEQLTMTTVDVAV 298
Db 239 AKDTFTLNGGPFAGFALAGVSLAHPDMESFIEARLNNAKQRTLKQI----- 288
Qy 299 RMAGDMGDFLDE-----PLEDILSTPEISHVFGDTKLGS-----AVPT 337
Db 289 ----RGRGFCPLPQVVLTYPFNLVFLVNDTNLLNEAPIASILKQETVVQABASVTVSPK 344
Qy 338 PVLIVQAVHDYLDVSDIDALADSYTAGGANVTYHRDLFSEHVSLLHPLSAPMTLRWLT 397
Db 345 PPRFIWHAIPDEIVPYQPAATYVKEQCAKANINPSPYIAEHLTAETIFGLVPSLWFIK 404
Qy 398 RFAGKPLTDHRVRTTWTPTIFNPMTYAGMARLAVIAAKVITGRKLSRR 444
Db 405 AFDG-----TTPKVICGTPIPALAGITTPSADQVGLGDLANQ 441

RESULT 8
US-08-360-758-2
; Sequence 2, Application US/08360758
; Patent No. 6074863
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Pathar, Shamkant A
; APPLICANT: Egel-Mitani, Michi
; APPLICANT: Borch, Kim
; APPLICANT: Clausen, Ib G
; APPLICANT: Hansen, Mogens T
; TITLE OF INVENTION: C. ANTARCTICA LIPASE AND LIPASE VARIANTS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESS: No. 6074863o No. 6074863disk of No. 6074863th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,758
; FILING DATE: 22-DEC-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK PCT/DK93/00225
; FILING DATE: 03-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3748.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 463 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

; MOLECULE TYPE: protein
US-08-360-758-2

Query Match.	16.4%;	Score	383;	DB	3;	Length	463;
Best Local Similarity	27.2%;	Pred.	No. 1.2e-27;				
Matches	127;	Conservative	64;	Mismatches	210;	Indels	66;
Gaps	10;						
QY	7	L A G A T G A E W I G R P H E L Q R K V R L L P S D P P F Y P P A G Y Q H A V P G T V L R S R D V E L A F M G L	66				
Db	12	L A A A T A A V - L A A P A A E T L D R R A A L P N Y D P P F Y T T P S N I G T F A K G V I Q S R K V - - - - -	63				
QY	67	I P O P V - - - - - T A T Q L L Y R T T M Y G N P E A T V T V I V P A E L A P G O T C P L L S Y Q C A I D A M S S	120				
Db	64	- P T D I G N A N N A S F O L Q Y R T I N T Q N E A V A D V A T W I P A K - P A S P P K I F S Y Q V Y E D A T A L	120				
QY	121	R C F P S Y A L R R R A K A L G S L T W - - E L M I S A A L A E G W A V S V P D H E G K L M G S P Y E P G Y R V	178				
Db	121	D C A P S Y L T G L D Q P N K V T A V L D T P I I G W A L Q Q Y Y V W S D H E G F K A A F I A G Y E E G W A I	180				
QY	179	L D G I R A A L N S E R V G L S P A T I G L W G Y S G G G L A S A W A A C E G Y A P D L D I V G A V L G S P V G D	238				
Db	181	L D G I R A L K N Y Q - - N L P S D K S V A L E G Y S G G A H A T W A T S L A E S Y A P E L I N V G A S H G G T P V S	238				
QY	239	L G H T F R R L N G T L A G L P A L V A A L Q H S Y P G L A R V I K E H A N D E G H Q L L E Q L T E M T T V D A V I	298				
Db	239	A K D T F T L N G G P F A G F A L A G V S G L S L A H P M O S E F I E A R L N A K G O R T L K Q I - - - - -	288				
QY	299	R M A G R D M G D F L D E - - - - - P L E D I L S T P E I S H V F G D T K L G S - - - - - A V P T	337				
Db	289	- - - - - R G R G F C L P Q V U L T Y P F L N V P S L V N D I N L N E A F I A L K O E T V V Q A E A S Y T S V S P K	344				
QY	338	P P V L I V Q A V H D Y L I D V S D I D A L A D S Y T A G G A N V T Y H R D L F S E H V S L H P L S A P M T L R W L T	397				
Db	345	F P R F I W H A I P D E I V P Y P A A T Y V K E Q C A K G A N I N F S P Y P I A E H L T A E I F G L V P S L W F I Q	404				
QY	398	R P A G K P L T D H R V R T T W T F I N P M T Y A G M A R L A V I A A K V I T G R K L S R R	444				
Db	405	A F D G - - - - - T T P K V I C G T P I P A I A G I T T P S A D Q V L G S D I A N Q	441				

```

RESULT 9
US-09-248-796A-16484
; Sequence 16484, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16484
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16484

Query Match      15.6%; Score 363; DB 4; Length 266;
Best Local Similarity 35.4%; Pred. No. 4.1e-26;
Matches 86; Conservative 46; Mismatches 103; Indels 8; Gaps 5;

Qy      30  PLLPSDDPFVFPFPPAGYOHAVPGTVLRSRSDVELAFM-GLIPQV-TATQLLYRTTNMVGNP 87
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      27  PTPKSIDPFYNAPGCFKNATVGDIQLQPKTKSITGCFVLNVQNSHQFVLVRSEDSEGNP 86

Qy      88  EATVTTVIVPAELAPGQTCPLLSYQCAIDAMSRCPFSYALRRRAKALGSLTQWELLMS 147
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      87  NVIVTTTVIEPVNADPSK--IASYQVSENAARADCAPSYALQFGSDVSTLTAQAEYLLA 143
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Qy 148 AALAEGWAVSVDPDHEGPKGLMGSPVEPGYRVLDGTIRAALNSERV-GLSPATPIGLWGYSG 206
      | : | | | | | : | | | | | : | | | | | : | | | | |
Db 144 PLLDKYVVVSPDYEGPKLTTVGKQSQOAVLNSIRASLKSCKITNIAEDAKVLMWGYSG 203
      | : | | | | | : | | | | | : | | | | | : | | | | |
Qy 207 GGLASAAWAAEACGEVAPDLDI--VCNAVIGSPVGDILGHTFRRLLNGTLLAGLPALVVAALOH 264
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 204 GSLTSGWAAALQPDYAPELSNLUGVALGGFITNVTTATVEATDDTI FAGIAANVLGGIAN 263
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 265 SYP 267
      ||
Db 264 EYP 266

```

```

RESULT 10
US-09-248-796A-16474
; Sequence 16474, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDATE
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16474
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16474

```

```

Query Match      15.1%; Score 351; DB 4; Length 285;
Best Local Similarity 39.1%; Pred. No. 6.2e-25;
Matches 81; Conservative 35; Mismatches 83; Indels 8; Gaps 5;

QY      30 PLLPSDDPFYFPFAGYOHAVPGTVLRSRDVELAFMG-LIPQV-TATQLLVRTTNMYGPN 87
Db      20 PTKPSSDPFFNPKPGFEKAAVGOLQSRTPKSTGRFAPLKQNSWQLLVRSDESGNP 79

QY      88 EATVTTTIVVPAELAPGTCPLLSYQCCAIDAMSSRCFFPSYALRRRAKALGSLTQMELLMIS 147
Db      80 NAIVTTTIVIEPNADPSK--IASYQVFEDAAKADCAPSYALQFGSDLTTFVTOAEMYLMA 136

QY      148 AALAEGWAVSVDPHEGPKGLWGSYPYEGYRVLVDCGIRAAALNSERV-GLSPATPTGLGWYSG 206
Db      137 PLLDQGYVVSPDYEGPKLFTTICKQSQAVLNSIRATLKSSKTNIKEDAKVVMWYSG 196

QY      207 GGLASMAAAEACGEYAPDLDIVGAVLG 233
Db      197 GSLASGWAALQPSYAPELS--SSLLG 221

RESULT 11
US-09-036-987A-14
; Sequence 14, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; TITLE OF INVENTION: Production
; NUMBER OF SEQUENCES: 39

```

;; CORRESPONDENCE ADDRESS:
;; ADDRESS: Dow Agrosciences LLC Patent Department
;; STREET: 9310 Zionsville Road
;; CITY: Indianapolis
;; STATE: Indiana
;; COUNTRY: USA
;; ZIP: 46268
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/036,987A
;; FILING DATE: 09-MAR-1998
;; CLASSIFICATION: 435
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Stuart, Donald R
;; REGISTRATION NUMBER: 28,479
;; REFERENCE/DOCKET NUMBER: 50,608
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (317)337-4816
;; TELEFAX: (317)337-4847
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 320 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-09-036-987A-14

Query Match 14.4%; Score 335.5; DB 3; Length 320;
Best Local Similarity 30.0%; Pred. No. 2.2e-23;
Matches 98; Conservative 49; Mismatches 147; Indels 33; Gaps 8;

Qy 87 PEATVTVIVPAELAPGQTCPLLSYQCAIDAMSSRCFPPSYALRRRAKALGSLTQWELLMI 146
Db 2 PNAVSGTVLVNIPWPREDRIITFAVGTGHLGSOVAPSYLLR-----TGTEPETELI 54

Qy 147 SAALAEWAVSPDHGPKGLWGSP-----YEGYRVLDGIRAAALNSERVCLSPATP 198
Db 55 AVALDRGWAVITDYE-----LGTGTHYTVGRAQGHAMLDAAARAAQRLPGSGLTDDCP 110

Qy 199 IGLWYSGGGLASAWAAEACGEYAPDLIVGAVLGSVPVGLDHTFRR---LNGTLIAGL 254
Db 111 VGWGYAQQGQASAFAGELHPTYAPELIRAAAGAVPIDLLDIHRNDGVFTGPVLGL 170

Qy 255 PALVVAALQHSYFGLARVIKEHANDEGRQLLEQLETTMTTVDVIRMAGRMDGDFLDPELE 314
Db 171 VGHAAA-----YPDLP--FDELLTEAGRTAVDQVRELGAPELVTRFLGRELSDFLD--TS 221

Qy 315 DILSTPEISHVFGDTKLG-SAVPTPPVLIQVAVHDYLDVSDIDALADSYTAGGANVTYH 373
Db 222 GLFEQPRWRARLAESVAGRNGGVPVPTLVYHSTDDEIVPFAGFGLRDSYRAAGTPVRWH 281

Qy 374 RDLFSEHVSUHLPLSAPMTLRLWLTDRFA 400
Db 282 PLUGLAHFFAALASSRVVSWFDEHFS 308

RESULT 12
US-09-370-700-14
; Sequence 14, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide

;; FILE REFERENCE: 50489 DIV1
;; CURRENT APPLICATION NUMBER: US/09/370,700
;; CURRENT FILING DATE: 1999-08-09
;; EARLIER APPLICATION NUMBER: US 09/36987
;; EARLIER FILING DATE: 1998-03-09
;; NUMBER OF SEQ ID NOS: 39
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 14
;; LENGTH: 320
;; TYPE: PRT
;; ORGANISM: Saccharopolyspora spinosa
;; US-09-370-700-14

Query Match 14.4%; Score 335.5; DB 3; Length 320;
Best Local Similarity 30.0%; Pred. No. 2.2e-23;
Matches 98; Conservative 49; Mismatches 147; Indels 33; Gaps 8;

Qy 87 PEATVTVIVPAELAPGQTCPLLSYQCAIDAMSSRCFPPSYALRRRAKALGSLTQWELLMI 146
Db 2 PNAVSGTVLVNIPWPREDRIITFAVGTGHLGSOVAPSYLLR-----TGTEPETELI 54

Qy 147 SAALAEWAVSPDHGPKGLWGSP-----YEGYRVLDGIRAAALNSERVCLSPATP 198
Db 55 AVALDRGWAVITDYE-----LGTGTHYTVGRAQGHAMLDAAARAAQRLPGSGLTDDCP 110

Qy 199 IGLWYSGGGLASAWAAEACGEYAPDLIVGAVLGSVPVGLDHTFRR---LNGTLIAGL 254
Db 111 VGWGYAQQGQASAFAGELHPTYAPELIRAAAGAVPIDLLDIHRNDGVFTGPVLGL 170

Qy 255 PALVVAALQHSYFGLARVIKEHANDEGRQLLEQLETTMTTVDVIRMAGRMDGDFLDPELE 314
Db 171 VGHAAA-----YPDLP--FDELLTEAGRTAVDQVRELGAPELVTRFLGRELSDFLD--TS 221

Qy 315 DILSTPEISHVFGDTKLG-SAVPTPPVLIQVAVHDYLDVSDIDALADSYTAGGANVTYH 373
Db 222 GLFEQPRWRARLAESVAGRNGGVPVPTLVYHSTDDEIVPFAGFGLRDSYRAAGTPVRWH 281

Qy 374 RDLFSEHVSUHLPLSAPMTLRLWLTDRFA 400
Db 282 PLUGLAHFFAALASSRVVSWFDEHFS 308

RESULT 13
US-09-603-207-14
; Sequence 14, Application US/09603207B
; Patent No. 6521406
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/603,207B
; CURRENT FILING DATE: 2000-06-23
; EARLIER APPLICATION NUMBER: 09/370,700
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 14
;; LENGTH: 320
;; TYPE: PRT
;; ORGANISM: Saccharopolyspora spinosa
;; US-09-603-207-14

Query Match 14.4%; Score 335.5; DB 4; Length 320;
Best Local Similarity 30.0%; Pred. No. 2.2e-23;
Matches 98; Conservative 49; Mismatches 147; Indels 33; Gaps 8;

Qy 87 PEATVTVIVPAELAPGQTCPLLSYQCAIDAMSSRCFPPSYALRRRAKALGSLTQWELLMI 146

Db 2 PNAVSTVLVNPWPBPRPIITFAVTHGLGSQVAPSYLLR-----TGTEPETELI 54
QY 147 SAALAEWAVSPDHGPKGLGSP-----YEPGYRVLGIRAAALNSERVGLSPATP 198
Db 55 AVALDRGMAVITDYEG---LGTPTHTYTVGRAQGHAMLDAAARAAQRLPGSLTTDCP 110
QY 199 IGLWYSGGLASAWAAACGAYAPLDIVGAVLSPVCDLGHFTFR-----LNGTLLAGL 254
Db 111 VGIWGYAQGGQSAFAGELHPTYPELRIRAAAGAVPDLDDIITHRNDGVFTGPFVLGL 170
QY 255 PALVVAALQHSYPLGARVLKHEANDGCOLLEQLEQTEMTTVDVAVRMAGRDMGDFLDEPLE 314
Db 171 VGHAAA-----YDLP--FDLLTEAGRTAVQVRELGAPELVTRFLGRELSDFLD--TS 221
QY 315 DILSTPEISHVFGDTKLG-SAVPTPVLIVQAVHVDYLDIVSDIDALADSYTAGGANVTYH 373
Db 222 GLFEQPRWRARLAESVAGRNGGPPVTLVYHSTDDEIVPFAFGERLDRSYRAAGTPVRMH 281
QY 374 ROLFSEHVSHPLSAPMTLRWLTDRFA 400
Db 282 PLSLAHFPAALASSRVVSWFDEHS 308

RESULT 14

US-09-602-787A-238
; Sequence 238, Application US/09602787A
; Patent No. 6696561
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Mark
; APPLICANT: Kr"ger, Burkhard
; APPLICANT: Sch"der, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; FILE REFERENCE: BGI-125CP
; CURRENT APPLICATION NUMBER: US/09/602,787A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: USSN 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931454.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931478.0
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931563.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932124.8
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932180.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932182.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932190.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932191.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932212.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932227.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932228.7
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932229.5

; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932927.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19940764.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940765.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940766.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940830.0
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940831.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940832.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940833.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941395.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942077.7
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942078.5
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942079.3
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 678
; SEQ ID NO 238
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-602-787A-238

Query Match 11.8%; Score 274.5; DB 4; Length 435;
Best Local Similarity 28.3%; Pred. No. 2.1e-17;
Matches 113; Conservative 57; Mismatches 184; Indels 45; Gaps 15;

QY 8 AGATGAEWIGRPP-----HEELQKVRPL--LPSDDPPY--FPPAGYQHAVP----GTVL 54
Db 14 AGSCDAWRTMSPLVQQGSEAVFRIMGLSRPRDRKPGFDDVPHFGAAVRVFGLKHGTLV 73
QY 55 RSRDVELAFMGLIPQVTTATOLLYRTTNMYGNPEATVTVIHPAELAPQOT--CPLLSYQ 112
Db 74 NAAPLKVLGARFENPASPSSYRFEYITGDSAG--RAITATGAVLFSTRPTWTTGPRAIANA 131
QY 113 CAIDAMSSRCPPSY--ALRRRA---KALGSLTQMELLMISAALAECAWVSPDHGPKGL 167
Db 132 PSTQGAQHCDPSHTCAIGLNAFYDKPFDATIAVELPVLWFLAHGLDVVFIDY----- 185
QY 168 WGSYPYEGYRV---LDGITRAA-----LNSERVGLSPATPIGLWYSGGLASAWAAARA 217
Db 186 ---PRDPATGVQYYCDSIAAAKSLDLDAVLASRQLGSPEAPLGLWGFSGGGATGNAQ- 241
QY 218 CGEYAPDLDIVCAVLGSPVGULGHTFRRLNGTLLAGLPALVVAALQHSYPLGARVLKEHA 277
Db 242 LQDYAPDRPKAAVVGAPVVDLFRVLDTVDGGLLTGVTIAYAIAGLVANSSSEFEMSVL 301
QY 278 NDEG-RQLEQLTEMTTVDVAVRMAGRDMG-DFLDEPLEDIL-STPEISHVFGDTKLGSA 334
Db 302 NERGVSVDLKNITSCAGGSLLASGYSRSSRGWTHOQTPLADILDDLPVVAERFKQKLGVR 361
QY 335 VTPPVLIVQAVHVDYLDIVSDIDALADSYTAGGANVTYH 373

Db 362 APEIPVLLWGSKNDVDVIPDPIRELDRSDWADKGTPLTWH 400

RESULT 15
US-09-602-777A-254
; Sequence 254, Application US/09602777A
; Patent No. 6831165
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zeider, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: BGI-128CP
; CURRENT APPLICATION NUMBER: US/09/602,777A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932930.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932933.8
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932935.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932973.7
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933002.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933003.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941390.8
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941391.6
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 254
; LENGTH: 435
; TYPE: PRT

; ORGANISM: Corynebacterium glutamicum
US-09-602-777A-254

Query Match 11.8%; Score 274.5; DB 4; Length 435;
Best Local Similarity 28.3%; Pred. No. 2.1e-17;
Matches 113; Conservative 57; Mismatches 184; Indels 45; Gaps 15;

Qy 8 AGATGAEWIGRPP-----HEELQKVRPL--LPSSDDPFY--FPPAGYQHAVP---GTVL 54
Db 14 AGCDAAWRWTMSPLVQOGSEAVFRIMGLSRPRDKGFDVPHFGAAVRVPGI.KHGTLV 73
Qy 55 RSRDVELAFMGLIPQPVTTATQLLYRTTMYGNPEATVTTIVPAELAPGQT--LPLLSYO 112
Db 74 NAAPLKVLGARGEENPASSYRFEVITGDSAG--RAITATGAVLFSTRPWTTPRPAIAMA 131
Qy 113 CAIDAMSSRCFPSY--ALRRRA---KALGSLTQWELLMISSAALAEGWAVSVDPDIEGPKGL 167
Db 132 PSTQGVAAQHCDCPSHTCAIGLNAFYDKPDAIIAIVELPVI.LWFLAHGI.DVVFI 185
Qy 168 WGSPEYEPGYRV---LDCIRAA-----LNSERVGLSPATPIGLWYSGGGLASAWAAEA 217
Db 186 ---PRDPATGVQYYCDSIAAKSLLDVAVLASRQLGSLPEAPLGWGF.SQGGATGWAQ- 241
Qy 218 CGEYAPDLDIVGAVLGSPVGDGLGHTFRRNGTLLAGLPALVVAALQHSYPGLARVVI KEHA 277
Db 242 LDQYAPDVRPKAAVVGAPPVDFRVLDTVDGGLLTGVIAVAIAGLAVNSSEMFEIAMS 301
Qy 278 NDEG-RQLLEQLTEMTTVDVIRNAGRDMG-DFLDEPLEDIL-STPEISHVFGDTKLGA 334
Db 302 NERGVSVDLVKNITSCAGGSLASCYSSSRGTHOGTPLADILDPLVVAEFKQKLG 361
Qy 335 VPTPPVLIVQAVHDYLDVSDIDALADSYTAGGANVTYH 373
Db 362 APEIPVLLWGSKNDVDVIPDPIRELDRSDWADKGTPLTWH 400

Search completed: October 15, 2005, 03:24:14
Job time : 38.6186 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:14:10 ; Search time 106.656 Seconds
(without alignments)
1742.576 Million cell updates/sec

Title: US-10-617-038-8
Perfect score: 2332
Sequence: 1 MVEPGLAGATGAEMIGRPP.....RLAVIAAKVITGRKLSRRPL 446

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications_AA.*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
 - 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
 - 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2332	100.0	446	15	US-10-617-038-8
2	621	26.6	446	14	Sequence 8, Appli
3	621	26.6	448	14	Sequence 3151, Ap
4	621	26.6	468	14	Sequence 3449, Ap
5	621	26.6	468	14	Sequence 8151, Ap
6	510.5	21.9	459	17	Sequence 8449, Ap
7	491.5	21.1	460	17	Sequence 105, App
8	455	19.5	465	17	Sequence 7020, Ap
9	439	18.8	468	14	Sequence 106, App
10	439	18.8	468	17	Sequence 7730, Ap
11	437.5	18.8	465	16	Sequence 104, App
					Sequence 2, Appli

12	437.5	18.8	471	16	US-10-681-636-4
13	416.5	17.9	466	14	US-10-032-585-7742
14	383	16.4	462	16	US-10-815-495-28
15	373.5	16.0	455	17	US-10-926-542-103
16	335.5	14.4	320	15	US-10-329-148A-14
17	274.5	11.8	435	15	US-10-627-476-238
18	272.5	11.7	408	9	US-09-738-626-5136
19	124	5.3	574	16	US-10-437-963-154545
20	120	5.1	631	18	US-10-450-763-39379
21	117	5.0	390	15	US-10-369-493-7771
22	116.5	5.0	753	14	US-10-080-170-246
23	116.5	5.0	753	16	US-10-080-170-246
24	116.5	5.0	753	16	US-10-468-356-246
25	113.5	4.9	433	16	US-10-425-115-238938
26	113.5	4.9	466	15	US-10-425-114-52082
27	113.5	4.9	572	15	US-10-425-114-52082
28	113.5	4.9	620	16	US-10-425-115-239157
29	112.5	4.8	391	15	US-10-344-738-91
30	112.5	4.8	7068	16	US-10-203-295-20
31	112.5	4.8	9477	16	US-10-203-295-37
32	111	4.8	494	15	US-10-369-493-17873
33	111	4.8	553	18	US-10-937-379-14
34	109.5	4.7	546	16	US-10-425-115-297522
35	109.5	4.7	706	16	US-10-437-963-155249
36	109	4.7	472	10	US-09-953-348-128
37	109	4.7	472	14	US-10-267-255-128
38	108.5	4.7	647	14	US-10-156-761-10968
39	108	4.6	389	17	US-10-732-923-4951
40	107	4.6	506	16	US-10-425-115-298724
41	107	4.6	1436	14	US-10-205-219-15
42	106.5	4.6	277	16	US-10-437-963-152420
43	106.5	4.6	424	14	US-10-156-761-7922
44	106.5	4.6	510	15	US-10-389-566-1551
45	106.5	4.6	1402	9	US-09-712-363-166

ALIGNMENTS

RESULT 1
US-10-617-038-8
; Sequence 8, Application US/10617038
; Publication No. US20040057963A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Peter
; APPLICANT: Rosenkrands, Ida
; APPLICANT: Stryhn, Anette
; TITLE OF INVENTION: Therapeutic TB Vaccine
; FILE REFERENCE: SS15AUSA
; CURRENT APPLICATION NUMBER: US/10/617,038
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: DK PA 2002 01098
; PRIOR FILING DATE: 2002-07-13
; PRIOR APPLICATION NUMBER: US 60/401,725
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-617-038-8

Query Match	100.0%	Score	2332;	DB	15;	Length	446;
Best Local Similarity	100.0%	Pred.	No. 2.1e-206;				
Matches	446;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	MVEPGLAGATGAEMIGRPPHEELQKVRPLPSDDPFYFPAGYOHAVPGTVLRSRDVE	60				
Db	1	MVEPGLAGATGAEMIGRPPHEELQKVRPLPSDDPFYFPAGYOHAVPGTVLRSRDVE	60				
QY	61	LAFMGLIPDPVTATQLLYRTTMYGNPEATVTVIVPAELAGQCPLLSYQCAIDAMSS	120				

Db 61 LAFMGLIPQVATQQLYRTTMYGNPEATVTVIAPAEAPGQTCPLLSYQCAIDAMSS 120
QY 121 RCFPSYALRRRAKALGUSTQWELLMSAALAEWAVSPDHEGPKGLWGSYPGGRVLD 180
Db 121 RCFPSYALRRRAKALGUSTQWELLMSAALAEWAVSPDHEGPKGLWGSYPGGRVLD 180
QY 181 GIRAALNSERVGLSPATPIGLWGYSGGLASAWAAEACGEYAPDLDIIVGAVLGSVPVGLG 240
Db 181 GIRAALNSERVGLSPATPIGLWGYSGGLASAWAAEACGEYAPDLDIIVGAVLGSVPVGLG 240
QY 241 HTFRLNGTLAAGLALVVAALQHSYPGLARVKEHNDGRLLEQLTEMTTVDVIRM 300
Db 241 HTFRLNGTLAAGLALVVAALQHSYPGLARVKEHNDGRLLEQLTEMTTVDVIRM 300
QY 301 AGRDMDGDFLEPLEDILSTPEISHVFGDTKLSAVPTPVLIVQAVHVDYLDIVSDIDALA 360
Db 301 AGRDMDGDFLEPLEDILSTPEISHVFGDTKLSAVPTPVLIVQAVHVDYLDIVSDIDALA 360
QY 361 DSYTAGGANVTYHRLDFSEHVSHPVLSAPMTLRWLTDRFAGKPLTDHVRVTTWTFIENPM 420
Db 361 DSYTAGGANVTYHRLDFSEHVSHPVLSAPMTLRWLTDRFAGKPLTDHVRVTTWTFIENPM 420
QY 421 TYAGMARLAVIAAKVITGRKLSRRPL 446
Db 421 TYAGMARLAVIAAKVITGRKLSRRPL 446

RESULT 2

US-10-128-714-3151
; Sequence 3151, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10128,714
; PRIOR FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3151
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-3151

Query Match 26.6%; Score 621; DB 14; Length 446;
Best Local Similarity 35.0%; Pred. No. 3e-48;
Matches 150; Conservative 74; Mismatches 173; Indels 32; Gaps 9;
QY 28 VRPLPSDDPPFPFAGYQHAVPGTVLRSRDV--ELAFMGLIPQVATQQLYRTTMY 84
Db 27 VRPL---EDPFYAPKGFESTVPGTILWRPNPNPISAFGAPINLAASYQLLYRSTDSF 83
QY 85 GNPEATVTVIAPAEAPGQTCPLLSYQCAIDAMSSRCFSPSYALRRRAKALGSLT----- 139
Db 84 GEPIAAASTILVPHN---ADNTKLLSFQAAEDANPNPCAPSAFQ-----LDSATDDELG 135

QY 140 ----QWELLMSAALAEWAVSPDHEGPKGLWGSYPGGRVLDGIRAAALNSERVGLSP 195
Db 136 LIMPQAEVLIIIAALDKGVWTVTPDHLGPNATFLANNLSGHVLDNIRALRSSAFSGISP 195
QY 196 ATPIGLWGYSGGLASAWAAEACGEYAPDLDIIVGAVLGSVPVGLDGLHTFRLNGTLAAGL 255
Db 196 KATITLWGYSGGSLASGLAAFLRASYPAPLNIAGNALGGTVPKIMPVFTVKNKGIYAGLI 255
QY 256 ALVVAALQHSYPGLARVKEHNDGRLLEQLTEMTTVDVIRMAGRDMDGDFLEPLED 315
Db 256 PAGQOGUSNYPFAIEKILYDHLVPAKKADPVTKNLCIVEDLLTYSIQDFVRYITD--AN 313
QY 316 ILSTPEISHVFGDTKLSAVPTPVLIVQAVHVDYLDIVSDIDALADSYTAGGANVTYHRD 375
Db 314 MKDPEVTRVLGENAMQHVDPDPLFVYKSNDDVSPVGGTDLALVSGYCAAGKVEYRD 373
QY 376 LFSHVSHPVLSAPMTLRWLTDRFAGKPL-TDHRVTRTWTTFIENPMYAGM-----ARL 428
Db 374 ELSNHTAMVIGVGNALLKDRMNGVPARAGCKTQTALTGLLDPTLAVLIGLIDLIKVL 433
QY 429 AVIAAKVIT 437
Db 434 ALJSAVGT 442

RESULT 3

US-10-128-714-3449
; Sequence 3449, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3449
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-3449

Query Match 26.6%; Score 621; DB 14; Length 468;
Best Local Similarity 35.0%; Pred. No. 3.2e-48;
Matches 150; Conservative 74; Mismatches 173; Indels 32; Gaps 9;
QY 28 VRPLPSDDPPFPFAGYQHAVPGTVLRSRDV--ELAFMGLIPQVATQQLYRTTMY 84
Db 49 VRPL---EDPFYAPKGFESTVPGTILWRPNPNPISAFGAPINLAASYQLLYRSTDSF 105
QY 85 GNPEATVTVIAPAEAPGQTCPLLSYQCAIDAMSSRCFSPSYALRRRAKALGSLT----- 139
Db 106 GEPIAAASTILVPHN---ADNTKLLSFQAAEDANPNPCAPSAFQ-----LDSATDDELG 157
QY 140 ----QWELLMSAALAEWAVSPDHEGPKGLWGSYPGGRVLDGIRAAALNSERVGLSP 195

Db 158 LIMPQELVLIITAAIDKGVVTVDPHLGNATFLANNLSCHVVLNINIRALRSSAFSGISP 217
Qy 196 ATPIGLWYSGGLASAWAAEACGEYAPDLDIVGAVLGSVPVGLDGHTRRLNGTLLAGLP 255
Db 218 KATITLWYSGGLASGLAAELRASVAPELNIAAGALGCTVPKIMPVNTVNGIYAGLL 277
Qy 256 ALVVAALQHSYGLARVKEHANDEGRQLLEQLTEMTTVDVIRWAGRDMGDFLDEPLED 315
Db 278 PAGMOGLSNEYPAIEKILYDHLVPAKADVFVKTKNLCEIVEDLLTYSFQDFYRIID--AN 335
Qy 316 ILSTPEISHVFGDTKLGSAVPTPPVLIIVQAVHDYLDIVSDIDALADSYTAGANVTYHRD 375
Db 336 MLKDEVTVLGENAMGQHVDPIDPLFVYKSTNDVSPVGDIDALVSGYCAAGCKVEYRD 395
Qy 376 LFEHVSLHPLSAPMTLRLWTRFAGKPL-TDHRVRTTWTPTIFNPMYAGM-----ARL 428
Db 396 ELSNHTMAVIGVNPALLMLKDRMGVPARAGCKTQTALTGLDPRTLAVLGIDLIKVL 455
Qy 429 AVIAAKVIT 437
Db 456 ALLSAPVGT 464

RESULT 4

US-10-128-714-8151
; Sequence 8151, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; PRIOR FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8151
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8151

Query Match 26.6%; Score 621; DB 14; Length 468;
Best Local Similarity 35.0%; Pred. No. 3.2e-48;
Matches 150; Conservative 74; Mismatches 173; Indels 32; Gaps 9;
Qy 28 VRPLPSDDPFYPPAGYQHAVPGTVLRSRDV--ELAFMGLIPQPVAT-QLLYRTTNNY 84
Db 49 VRPL---EDPFYSAPKGFESTVPGTILRWNPENPISAFGFAPINLAASYQLLYRSTDSF 105
Qy 85 GNPEATVTTVIIPAEALAPGCTCLLSYQCAIDAMSSRCFPPSVALRRKALGSLT----- 139
Db 106 GEPIAAASTILVPHN---ADNTKLLSFQAAEDAAANPCAPSAFO-----LDSATDDELG 157
Qy 140 ----QNELMISAALAEWAVSPDHEGPKLWGSYPGPGYRVLDCIRALNSERVGLSP 195
Db 158 LIMPQELVLIITAAIDKGVVTVDPHLGNATFLANNLSCHVVLNINIRALRSSAFSGISP 217

Qy 196 ATPIGLWYSGGLASAWAAEACGEYAPDLDIVGAVLGSVPVGLDGHTRRLNGTLLAGLP 255
Db 218 KATITLWYSGGLASGLAAELRASVAPELNIAAGALGCTVPKIMPVNTVNGIYAGLL 277
Qy 256 ALVVAALQHSYGLARVKEHANDEGRQLLEQLTEMTTVDVIRWAGRDMGDFLDEPLED 315
Db 278 PAGMOGLSNEYPAIEKILYDHLVPAKADVFVKTKNLCEIVEDLLTYSFQDFYRIID--AN 335
Qy 316 ILSTPEISHVFGDTKLGSAVPTPPVLIIVQAVHDYLDIVSDIDALADSYTAGANVTYHRD 375
Db 336 MLKDEVTVLGENAMGQHVDPIDPLFVYKSTNDVSPVGDIDALVSGYCAAGCKVEYRD 395
Qy 376 LFEHVSLHPLSAPMTLRLWTRFAGKPL-TDHRVRTTWTPTIFNPMYAGM-----ARL 428
Db 396 ELSNHTMAVIGVNPALLMLKDRMGVPARAGCKTQTALTGLDPRTLAVLGIDLIKVL 455
Qy 429 AVIAAKVIT 437
Db 456 ALLSAPVGT 464

RESULT 5

US-10-128-714-8449
; Sequence 8449, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; PRIOR FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8449
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8449

Query Match 26.6%; Score 621; DB 14; Length 468;
Best Local Similarity 35.0%; Pred. No. 3.2e-48;
Matches 150; Conservative 74; Mismatches 173; Indels 32; Gaps 9;
Qy 28 VRPLPSDDPFYPPAGYQHAVPGTVLRSRDV--ELAFMGLIPQPVAT-QLLYRTTNNY 84
Db 49 VRPL---EDPFYSAPKGFESTVPGTILRWNPENPISAFGFAPINLAASYQLLYRSTDSF 105
Qy 85 GNPEATVTTVIIPAEALAPGCTCLLSYQCAIDAMSSRCFPPSVALRRKALGSLT----- 139
Db 106 GEPIAAASTILVPHN---ADNTKLLSFQAAEDAAANPCAPSAFO-----LDSATDDELG 157
Qy 140 ----QNELMISAALAEWAVSPDHEGPKLWGSYPGPGYRVLDCIRALNSERVGLSP 195
Db 158 LIMPQELVLIITAAIDKGVVTVDPHLGNATFLANNLSCHVVLNINIRALRSSAFSGISP 217
Qy 196 ATPIGLWYSGGLASAWAAEACGEYAPDLDIVGAVLGSVPVGLDGHTRRLNGTLLAGLP 255

Db 218 KATITLWYSGSLASGLAAELRASYPALNIAAGALGGTVKIMPVNTVKNKIYAGLL 277
QY 256 ALVVAALQHSYPLARVKEHANEDEGRQLLEQLTEMVTVDAVIRWAGDMGDFLDEPLED 315
Db 278 PAGMOGLSNEYPAIEKILYDLHVPAAKADPFVKTNKLCIVEDLLTVSFQDFYRYITD--AN 335
QY 316 ILSTPEISHVFGDTKLGSAVPTPPVLIVQAVHDYLDVSDIDALADSYTAGGANVTVHRD 375
Db 336 MLKDEPVTVLGENAMGQHVDPIDLPVYKSTNDVSPVGDTDALVSGVCAAGKGVYRDR 395
QY 376 LFSEHVSUHLPSAPMTLRLWLTDRFAGKPL--TDHVRVTTWPTTFINPMTYAGM-----ARL 428
Db 396 ELSNHTAMVIGVNPALLWLKDRMNGVPARACKTQTALTGLDPRTLAVLIDILIKVLL 455
QY 429 AVIAAKVIT 437
Db 456 ALLSAPVGT 464

RESULT 6

US-10-926-542-105

; Sequence 105, Application US/10926542

; Publication No. US20050053989A1

; GENERAL INFORMATION:

; APPLICANT: LABAN, ABRAHAM

; APPLICANT: SHARON, GIL

; TITLE OF INVENTION: LIBRARIES OF RECOMBINANT PROTEINS

; FILE REFERENCE: 178.002

; CURRENT APPLICATION NUMBER: US/10/926.542

; CURRENT FILING DATE: 2004-08-26

; PRIOR APPLICATION NUMBER: US 60/497,924

; PRIOR FILING DATE: 2003-08-27

; NUMBER OF SEQ ID NOS: 127

; SOFTWARE: PatentIn Ver. 3.2

; SEQ ID NO 105

; LENGTH: 459

; TYPE: PRT

; ORGANISM: Candida albicans

US-10-926-542-105

Query Match 21.9%; Score 510.5; DB 17; Length 459;

Best Local Similarity 31.7%; Pred. No. 5.1e-38;

Matches 132; Conservative 77; Mismatches 166; Indels 41; Gaps 12;

QY 28 VPLLPSPDDPFYFPAGYQHAVPGTVLRSRDVELAFMG-LIPQPV-TATQLLYRTTNNMYGNP 77
Db 17 ILPTKPSNDPFYNPAGFEKAAVGDILQSRKT-----PKPITGVFVVKIQQNSWQLL 68

QY 78 YRTTNMYGNPEATVTVTVPAELAPGTCPLLSYQCAIDAMSSRCFSPSYALRRRAKALGS 137
Db 69 VRSEDSFGNPNVITVTNVEPNADPSK---LASIQVFEDSAKADCAPSYALQFGSDVTTI 125

QY 138 LTQWELLMSAALAEWAVSPDHGPKGLNGSPYEPGYRVLGDIRAALNSERV-GLSPA 196
Db 126 ATQVETYLAPLLDQGYVYVSPDYEGPKSTETVKGQSQAVLNSTRALKSKGITNAEN 185

QY 197 TPIGLWYSGGSLASAWAAEACGEYAPDL--DIVGAVLGSVPGDLGHTFRLNGTLLAGL 254
Db 186 AKVVMWYSGGSLASGWAALQPNVAPELGNGLLGAALGGFTVNTATAEATDGTGVFAGI 245

QY 255 PALVVAALQHSYPLARVKEHANEDEGRQ-LLEQLTEMVTVDAVIRWAGDMGDFLDEPL 313
Db 246 MANALGGVANEPKQILQ---NDTDKQSVDFQDNHCLADGVINYIGKHFSLGSKNIF 302

QY 314 E---DILTSTPEISHVFGDTKL---GSAVPTPPVLIVQAVHDYLDVSDIDALADSYTAG 366
Db 303 KSGWNILKNPTISKIVEDNGLVYQKQLVPKIPILYHGAIDQIVPVNVKKTQNWCDAG 362

QY 367 GANVTVHRDLFSEHVSUHLPSAPMTLRLWLTDRFAGKPLT---DHRVRTT---WPTI 416

Db 363 IASLEFSESDATNGHITETIVGAPVALTWIIRFNGKQTVSGCQHVKRTSNFEYFNI 418

RESULT 7

US-10-741-849-7020

; Sequence 7020, Application US/10741849

; Publication No. US20050019931A1

; GENERAL INFORMATION:

; APPLICANT: Roemer, Terry

; APPLICANT: Jiang, Bo

; APPLICANT: Boone, Charles

; APPLICANT: Bussey, Howard

; TITLE OF INVENTION: Nucleic Acids Encoding Anti-fungal Drug Targets and Methods of

; FILE REFERENCE: 10182-023-999

; CURRENT APPLICATION NUMBER: US/10/741.849

; CURRENT FILING DATE: 2003-12-19

; PRIOR APPLICATION NUMBER: US 60/434,832

; PRIOR FILING DATE: 2002-12-19

; NUMBER OF SEQ ID NOS: 8000

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 7020

; LENGTH: 460

; TYPE: PRT

; ORGANISM: Candida albicans

US-10-741-849-7020

Query Match 21.1%; Score 491.5; DB 17; Length 460;

Best Local Similarity 30.5%; Pred. No. 2.9e-36;

Matches 120; Conservative 78; Mismatches 182; Indels 33; Gaps 10;

QY 30 PLLPSDDPFYFPAGYQHAVPGTVLRSRDVELAFMG-LIPQPV-TATQLLYRTTNNMYGNP 87
Db 19 PTKPSSDPFYNPAGFEKAAVGDILQSRTPKSTGRFAPLKIQNSQQLLVRSDDSPGNP 78

QY 88 EATVTVTVPAELAPGTCPLLSYQCAIDAMSSRCFSPSYALRRRAKALGSITQWELLMS 147
Db 79 NAIVTVTVIEPNADPSK---IASYQVFEDAAKADCAPSYALQFGSDJTTFTVTAEMVYLM 135

QY 148 AALAEWAVSPDHGPKGLNGSPYEPGYRVLGDIRAALNSERV-GLSPATPIGLWYSG 206
Db 136 PLDDQGYVYVSPDYEGPKSTFTICKQSQAVLNSIRATLKSSKITNIKEDAKVVMWYSG 195

QY 207 GGLASAWAAEACGEYAPDL--DIVGAVLGSVPGDLGHTFRLNGTLLAGLPALVVAALQH 264
Db 196 GSLASGWAALQPSYAPELSSLLGAALGGFTVNTATQAADGTVFAGIVANALGGVAN 255

QY 265 SYPCGLARVKEHANEDEGRQLLEQLTEMVTVDAVIRWAGDMGDFLDEPL----- 313
Db 256 EYEPKSILO--SDTDKKSVDPEFDSHCLADGVI-----DYINTSFLTGDNKIFKTG 305

QY 314 EDILTSTPEISHVFGDTKL---GSAVPTPPVLIVQAVHDYLDVSDIDALADSYTAGG-AN 369
Db 306 WDLKSPITAKIVEDNGLVYQKQLVPKIPFVYHGSIDQIVPVNVKKTQYONWCEGGISS 365

QY 370 VTYHRDLFSEHVSUHLPSAPMTLRLWLTDRFAGK 402
Db 366 LEFAEDGTNGHLETTVVGAPAAALTWIIDRFENGK 398

RESULT 8

US-10-926-542-106

; Sequence 106, Application US/10926542

; Publication No. US20050053989A1

; GENERAL INFORMATION:

; APPLICANT: LABAN, ABRAHAM

; APPLICANT: SHARON, GIL

; TITLE OF INVENTION: LIBRARIES OF RECOMBINANT PROTEINS

; FILE REFERENCE: 178.002

; CURRENT APPLICATION NUMBER: US/10/926.542

; CURRENT FILING DATE: 2004-08-26

; PRIOR APPLICATION NUMBER: US 60/497,924

; PRIOR FILING DATE: 2003-08-27

; NUMBER OF SEQ ID NOS: 127

; SOFTWARE: PatentIn Ver. 3.2

; SEQ ID NO 106


```
RESULT 11
US-10-681-636-2
; Sequence 2, Application US/10681636
; Publication No. US20040142441A1
; GENERAL INFORMATION:
; APPLICANT: Weiss, Albrecht
; APPLICANT: Bigey, Frederic
; APPLICANT: Dubreucq, Eric
; APPLICANT: Moulin, Guy
; TITLE OF INVENTION: Enzymes with Lipase/Acyltransferase
; TITLE OF INVENTION: Activity, Nucleic Acids Encoding the Same and Methods of Use
; FILE REFERENCE: 2376-P03297US01
; CURRENT APPLICATION NUMBER: US/10/681,636
; PRIOR FILING DATE: 2003-10-08
; PRIOR FILING DATE: 60/416,987
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Candida parapsilosis
US-10-681-636-2

Query Match      18.8%; Score 437.5; DB 16; Length 465;
Best Local Similarity 30.0%; Pred. No. 2.9e-31;
Matches 127; Conservative 70; Mismatches 159; Indels 67; Gaps 12;

QY 30 PLLPSDDPFYPPAGYQHAVPGTVLRSRDVELAFMGLIPQPV-----ATQLLYR 79
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21 PKPSQDDFYTPQGYEAQPLGSLKTRNV-----PNLTNVFTPVKVQNAQLLVR 72
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 80 TNNYGNPEATVTTVIVPAELAPGQTCPLLSYQCAIDAMSSRCPPSYALRRRAKALGSLT 139
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 73 SEDTFGNPNAIVTTIIQPFNAKKDK--LVSYQTFEDSGKLDCAPSYAIQYGSIDISTLT 129
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 140 QMELLMISAALAEGWAVSPDHGPKLWGPYRGYVLDGIRAAALNSERV-GLSPATP 198
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 130 QEMYIYSALLDQGYVYVTPDYEGPKSTFTVGLQSGRATLNSRLATKSGNLTGVSSDAE 189
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 199 IGLWYSGGGLASAWAAEACGEYAPDL--DIVGAVLGSVPDGLGHTFRRNLGTLLAGLPA 256
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 190 TLLWYSGGSLASGAAAIQKEYAPELSKNLLGALGGFVNTITATAEAVDSGPFAGIIS 249
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 257 LVVAALQHSYPLGLARVIKEHANDEGRQLLEQLTMTTV-----DAVIMAGR-- 303
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 250 NALAGIGNEYP-----DFKNYLLKKVSPLLSITYRLGNTHCCLDGGIAYFGKSF 298
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 304 -----DMGDFLD-EPLEDILTPEISHVFGDTKLGSVPTPPVLIQVAVHDYLDV 353
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 299 FSRIRYFPDGMWLVNQEPKIKTILQDNLGLVYQPKD-----LTPQIPLFYHGTLDIVPI 353
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 354 SDIDALADSYTAGG-ANVTYHRDLFSEHVSHPLSAPMTLRWLTDRFAGKPLTD---HRV 409
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 354 VNSRKTFCQCDWGLKSGEYNEDLTNGHITESIVGAPAAALTWIINRFGPPVDCQHN 413
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 410 RTT 412
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 414 RAS 416

RESULT 12
US-10-681-636-4
; Sequence 4, Application US/10681636
; Publication No. US20040142441A1
; GENERAL INFORMATION:
; APPLICANT: Weiss, Albrecht
; APPLICANT: Bigey, Frederic
; APPLICANT: Dubreucq, Eric
; APPLICANT: Moulin, Guy
; TITLE OF INVENTION: Enzymes with Lipase/Acyltransferase
; TITLE OF INVENTION: Activity, Nucleic Acids Encoding the Same and Methods of Use
```

```
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 2376-P03297US01
; CURRENT APPLICATION NUMBER: US/10/681,636
; PRIOR FILING DATE: 2003-10-08
; PRIOR FILING DATE: 60/416,987
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Candida parapsilosis
US-10-681-636-4

Query Match      18.8%; Score 437.5; DB 16; Length 471;
Best Local Similarity 30.0%; Pred. No. 3e-31;
Matches 127; Conservative 70; Mismatches 159; Indels 67; Gaps 12;

QY 30 PLLPSDDPFYPPAGYQHAVPGTVLRSRDVELAFMGLIPQPV-----ATQLLYR 79
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21 PKPSQDDFYTPQGYEAQPLGSLKTRNV-----PNLTNVFTPVKVQNAQLLVR 72
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 80 TNNYGNPEATVTTVIVPAELAPGQTCPLLSYQCAIDAMSSRCPPSYALRRRAKALGSLT 139
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 73 SEDTFGNPNAIVTTIIQPFNAKKDK--LVSYQTFEDSGKLDCAPSYAIQYGSIDISTLT 129
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 140 QMELLMISAALAEGWAVSPDHGPKLWGPYRGYVLDGIRAAALNSERV-GLSPATP 198
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 130 QEMYIYSALLDQGYVYVTPDYEGPKSTFTVGLQSGRATLNSRLATKSGNLTGVSSDAE 189
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 199 IGLWYSGGGLASAWAAEACGEYAPDL--DIVGAVLGSVPDGLGHTFRRNLGTLLAGLPA 256
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 190 TLLWYSGGSLASGAAAIQKEYAPELSKNLLGALGGFVNTITATAEAVDSGPFAGIIS 249
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 257 LVVAALQHSYPLGLARVIKEHANDEGRQLLEQLTMTTV-----DAVIMAGR-- 303
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 250 NALAGIGNEYP-----DFKNYLLKKVSPLLSITYRLGNTHCCLDGGIAYFGKSF 298
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 304 -----DMGDFLD-EPLEDILTPEISHVFGDTKLGSVPTPPVLIQVAVHDYLDV 353
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 299 FSRIRYFPDGMWLVNQEPKIKTILQDNLGLVYQPKD-----LTPQIPLFYHGTLDIVPI 353
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 354 SDIDALADSYTAGG-ANVTYHRDLFSEHVSHPLSAPMTLRWLTDRFAGKPLTD---HRV 409
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 354 VNSRKTFCQCDWGLKSGEYNEDLTNGHITESIVGAPAAALTWIINRFGPPVDCQHN 413
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 410 RTT 412
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 414 RAS 416

RESULT 13
US-10-032-585-7742
; Sequence 7742, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7742
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7742

Query Match      17.9%; Score 416.5; DB 14; Length 466;
```


Qy 410 RTTPTIENPTYAGMARLAVIAAKVITCRKLSRR 444
Db 402 --TTPKVICGTPIPAIAIGITTPSADOVLGSDIANQ 434

Search completed: October 15, 2005, 04:14:59
Job time : 107.656 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:59:30 ; Search time 24.8864 Seconds
(without alignments)
1724.342 Million cell updates/sec

Title: US-10-617-038-8
Perfect score: 2332
Sequence: 1 MVEPGLAGATGAEWIGRPP.....RLAVIAKVITGRKLSRRRL 446

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: +
1: PIR1: +
2: PIR2: +
3: PIR3: +
4: PIR4: +

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2332	100.0	446	2	B70543
2	492.5	21.1	463	2	hypothetical prote
3	391	16.3	458	2	triacylglycerol li
4	257	11.0	132	2	Lep1170 Fl 46 pro
5	116.5	5.0	577	2	hypothetical prote
6	116.5	5.0	753	2	probable DNA-bindi
7	112.5	4.8	603	2	gamma-glutamyltran
8	111.5	4.8	575	2	penicillin-binding
9	106.5	4.6	752	2	probable ctpB prot
10	106.5	4.6	1402	2	probable polyketid
11	106	4.5	792	2	phenylalanyl-tRNA
12	106	4.5	1436	2	CJ5290
13	106	4.5	2611	2	protein-tyrosine-p
14	105	4.5	728	2	actinomycin synthe
15	105	4.5	1058	2	ankyrin-like prote
16	105	4.5	3519	2	polyketide synthas
17	104.5	4.5	420	2	hypothetical prote
18	104.5	4.5	1827	2	probable polyketid
19	104	4.5	474	1	cyae protein - Bor
20	104	4.5	761	2	probable ctpA prot
21	104	4.5	1047	2	GTPase-activating
22	104	4.5	1892	2	carbamoyl-phosphat
23	104	4.5	1997	2	virginiamycin S sy
24	103.5	4.4	351	2	threonine synthase
25	103.5	4.4	883	2	probable ftsK - My
26	103	4.4	291	2	short-chain dehydr
27	103	4.4	291	2	probable short-cha
28	103	4.4	993	2	glycine decarboxyl
29	102.5	4.4	566	2	probable two-compo

ALIGNMENTS

RESULT 1

B70543
hypothetical protein Rv1592c - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: B70543
R: Cole, S.T.; Davies, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A: Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A: Reference number: A70500; MUID: 98295987; PMID: 9634230
A: Accession: B70543
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-446 <COL>
A: Cross-references: UNIPROT: O06598; GB: Z95586; GB: AL123456; NID: g3261785; PIDN: CAB09077.1
A: Experimental source: strain H37Rv
C: Geneticks:
A: Gene: Rv1592c
C: Superfamily: Mycobacterium tuberculosis hypothetical protein Rv1592c

Query Match	Best Local Similarity	Score	DB 2	Length	446;
Matches	446;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;		
Qy	1	MVEPGLAGATGAEWIGRPPHEELQKVRPLLP	SDDDPFYFPDPAGYQHAVPGTVLRSRDVE	60	
Db	1	MVEPGLAGATGAEWIGRPPHEELQKVRPLLP	SDDDPFYFPDPAGYQHAVPGTVLRSRDVE	60	
Qy	61	LAFMGLIPQPVTTATOLLYRTTNMYGNPEATV	TVTPAEALAPGQTCPLLSYOCADAMSS	120	
Db	61	LAFMGLIPQPVTTATOLLYRTTNMYGNPEATV	TVTPAEALAPGQTCPLLSYOCADAMSS	120	
Qy	121	RCFPSYALRRRAKALGSLTQMELLMISAALAE	GWAVSPDHEGPKGLMGSPYEPGRVILD	180	
Db	121	RCFPSYALRRRAKALGSLTQMELLMISAALAE	GWAVSPDHEGPKGLMGSPYEPGRVILD	180	
Qy	181	GIRAAINSERVGLSPATPIGLWYSGGGGLAS	AWAAEACGEYAPDLDIVAGVLGSPVGDILG	240	
Db	181	GIRAAINSERVGLSPATPIGLWYSGGGGLAS	AWAAEACGEYAPDLDIVAGVLGSPVGDILG	240	
Qy	241	HTFERLNGTLTLAGLPALVVAALOHSHYPGL	ARVIRKHANDEGROLLEQLEMTTVDVAVIRM	300	
Db	241	HTFERLNGTLTLAGLPALVVAALOHSHYPGL	ARVIRKHANDEGROLLEQLEMTTVDVAVIRM	300	
Qy	301	AGRDMDGLDLEPDLSTPEISHVFGDTKLGSA	VPTPPVLIVQAVHYDLIDVSDIDALA	360	
Db	301	AGRDMDGLDLEPDLSTPEISHVFGDTKLGSA	VPTPPVLIVQAVHYDLIDVSDIDALA	360	
Qy	361	DSYTAGGANVTYVHRDLFSEHVSLSLPSAPMT	LRWLTDRFAGKPLTDHRVRRVTWPTIFNPM	420	

Db 361 DSYTAGCANVTYHRDLFSEHVSLSPLSAPMLRLWLTDRFAGKPLTDHRVTPTIFNEM 420
QY 421 TVAGMARLAVIAAKVITGRKLSRRPL 446
Db 421 TVAGMARLAVIAAKVITGRKLSRRPL 446

RESULT 2

T18225
hypothetical protein - yeast (*Candida albicans*)
C:Species: *Candida albicans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18225
R:Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, November 1998
A:Reference number: Z18831
A:Accession: T18225
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-463 <BAR>
A:Cross-references: UNIPROT:O94015; EMBL:AL033396; NID:e1340089; PID:e1340098; PIDN:CAA2
C:Genetics:
A:Note: C35A5.09
C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv1592c

Query Match	21.1%	Score 492.5	DB 2	Length 463
Best Local Similarity	33.2%	Pred. No. 1.2e-28		
Matches 128	Conservative	70	Mismatches 169	Indels 19
				Gaps 11

Qy 30 LLLLSDDDFYFPAGYQHAVPGTVLRSRDVELAFM-GLIPQPV-TATQLLYRTTNMYGNP 87
| | | | | : : : : : : : : : : : : : : : : : :
Dd 19 PTKTSSDFFYNPKGFENAAVGDIQSRA TPKSTGGFTPLKINQSWOLLVRSSEFSFGNP 78
| | | | | : : : : : : : : : : : : : : : : : :

QY 88 EATVTTVIPAEIAPGQTCPLLSYCCAIDAMSSRCFPVSVALRRRAKALGSLTQMELLMIS 147
||||| | : : : ||| : ||||| : ||| : ::
Db 79 NVIVTTVIPNVNADPSK---IASOVFEAAAKACAPSVALOFGSDITTVTVOAERWIMA 135

Qy

148	AALAEGWAVSPDHEGPKLWGSPEYGRVLDGIRAAINSERV-GLSPATPIGIWGYSG	206
: : :	: : :	: : :

Db

136	PILDGYYVVSPNYEGPKSTFTTCOSGOAIVNSTRAALKSGKTTNTKGDAKVMMWYGSG	195
: : :	: : :	: : :

Qy 207 GGIASAWAAEACGEYAPDL--DIVGAVLGSPVGD LGHTFRRNLGTLLAGLIPALVVAA LQH 264
| | | | | | | : : : : : : : : : : : : : : :
Db 196 GSIIASGWAAALQPSYAPEIGNNILGAALGGFVTNITATCAATDCGFVFAGSYVANALGC VAN 255

Qy	265	SYPGLARVIKEHANDEGRO-LLEQLTENTTVDAVIRMGDRM--GDF-LDEPLEDILSTP	320
		: : : : : : : : : : : : : : : : : : : : : : : :	
Db	256	EYSEFKOVLQ---NDTDQOSTEAFENSNCI IDGVINVTCTSETTCDDKIVETKTQNTI KNP	313

Qy 321 EISHVFGDTK---GSAVPTPPVLIVQAVHDYLI DVSDIDALASY-TAGGANVTYHRDL 376

Qy 377 FSEHVSLHPLSAPWTLRWLTDRFAGK 402
:
:
:
Dk 378 GCVLSTENINCAIRANVTHTITDFEYK 398
:
:
:

RESULT 3

triacylglycerol lipase (EC 3.1.1.3) - *Kurtzmanomyces* sp. I-11
 N:Alternate names: lipase; triacylglycerol acylhydrolase; tributyrase; triglyceride lipase
 C:Species: *Kurtzmanomyces* sp. I-11
 C:Date: 09-Dec-2002 #sequence_revision 09-Dec-2002 #text_change 09-Jul-2004
 C:Accession: JG7868; PC7198
 A:Kakugawa, K.; Shobayashi, M.; Suzuki, O.; Miyakawa, T.
 Biosci. Biotechnol. Biochem. 66, 1328-1336, 2002
 A:Title: Cloning, characterization, and expression of cDNA encoding a lipase from *Kurtzmanomyces*
 A:Reference number: JG7868; MIMD:22152189; PMID:12162555
 A:Accession: JG7868
 A:Molecule type: mRNA
 A:Residues: 1-458 <RAK>

A;Cross-references: UNIPROT:Q8N1P2; DDBJ:AB073866
A;Accession: PC7198
A;Molecule type: protein
A;Residues: 27-56;133-152;188-207;301-317;340-357 >KA2>
C;Comments: This enzyme, which is a thermostable, acidophilic and position-independent of glycerolides. Thus, this enzyme is useful as a catalyst in industry such as C;Genetics:
A;Gene: lip1
C;Keywords: carboxylic ester hydrolase

Query Match	16.3%	Score 381;	DB 2;	Length 458;
Best Local Similarity	27.0%;	Pred. No. 2e-20;		
Matches 118: Conservative	69;	Mismatches 188;	Indels	62; Gaps 12

Qy	33	PSDDPFYPPAGYQHVA	PGTVLSRDELAFMGLI	PPQV-----	TATQLLYRTTMYGN	86
Db	32	PNEDPPYSTPSNIET	FANGOIIGSRKV-----	PTDIGNSNAAS	YOLSYRTNTOED	83

Qy	87	PEATVTTVVIPAEAPGGTCPLLSYQCAIDAMSSRCFPSYA	---LRRRAKALGSLTOME	142
Dh	84	AVANVATVTPAK--PSSPRTTYQVYFSDTGLDCAPSYSVTCGVDBNKATAVI	DTB-140	

[illegible]

QY	262
QB	257

[illegible]

QY 310 DEPLEDILSTPETSHVFGDTKLGSAVTPP - PVLIVQAVHDYLLDVSDIDALADSYTAGG 367

QY 368 ANVTYHRDLFSEHVSLHPLSAPMTLRWLTDTRFAGKPLTDHVRVTTPTIENBMTYAGMAR 427

Qy 428 LAVIAAKVITGRKLSRR 444

RESULT 4

Lepb1170_F1_46 protein - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

R:Smith, D.R.; Robison, K.
submitted to the EMBL Data Library, November 1993
A:Description: *Mycobacterium leprae* cosmid B1170.

A;Accession: S72713
A;Status: preliminary
A;Molecule type: DNA

A;Cross-references: UNIPROT:Q49629; EMBL:U00010; NID:q466780; PIDN:AAA17077.1; PID:q46680

Query Match 11.0%; Score 257; DB 2; Length 132;
Best Local Similarity 56.3%; Pred. No. 5.4e-12;
Matches 49; Conservative 14; Mismatches 24; Indels

Qy	262	LQHSY	PGLAR	VIKE	HANDE	GRQLE	QLLEQ	LTMTT	VDVIR	MAGRO	MGDF	LDEPLE	LILST	PE	321
Db	46	LSHTY	PALDK	VIKE	HANDE	GRAL	PTRLK	OMTT	VGAVI	KMAGR	FSFY	IHKPLA	ETLSR	PT	105

Qy 322 ISHVFGDTKLGSAPVTPPVLIVQAVHD 348

: : : : : : : : : : :

Db 106 VAHVNNIKLSSGVPNRQLIIVQAIHD 132

RESULT 5
T45333
hypothetical protein MLCB57.01c [imported] - Mycobacterium leprae (fragment)
C:Species: Mycobacterium leprae
C:Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jul-2004
C:Accession: T45333
R:Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z16918
A:Accession: T45333
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-577 <PAR>
A:Cross-references: UNIPROT:O33047; EMBL:Z99494; PIDN:CAB16641.1
A:Experimental source: cosmid B57
C:Genetics:
C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0862c

Query Match 5.0%; Score 116.5; DB 2; Length 577;
Best Local Similarity 25.0%; Pred. No. 0.87;
Matches 124; Conservative 46; Mismatches 209; Indels 117; Gaps 24;

QY 4 PGNLA----GATG---AEWIGRPPHEELQKVRPLLPSPDDPFYFPAGYQHAVPGTVLRS 56
Db 132 PQQVTLIEDISSTGEQIAELIAR--LSQTQRDVLQKLEGSPL----GRTRDAAPGAPPDR 185

QY 57 RDVELAFMGLIPQPVATATQLLYRTNNY-----GNPEAT-----VTTVIIVP-----AELAP 102
Db 186 PVPQLLAMGLLRIRIDADTVILPRRVQVLRGEQSGPTQLTQPVVVSVTTTPNDADAEAG 245

QY 103 GOTCPILLSYQCAIDAMSSRCFPSYALR-----RRAKALG-SLTOMELLMISAALA 151
Db 246 AVIEALHELVDVLETLGSA--PVYELRNGGLGVREFKRLAKATGINEPRGLLLEVTAA 303

QY 152 EGVAVSPDHEGPKLWGSPYEGYRVLDIGIRAAALNSERVGLSPATPIGLWYSGGGLAS 211
Db 304 GLIASGIPDPEPATG--DSPYNAPTATDRFTAMPAPAEKRWHL-----LAS 346

QY 212 ANAAEAC-----GEYAPDDIDVGAVLGSPVGDIGHTFRRLNGTLLAGLPALV-----V 259
Db 347 TWLDLQCRPALIGSRGPDAKSYGALSNSLYSTAAPLDRRLLLGMLAELPAGVGVEAAEAS 406

QY 260 AALQHSYPGLAR-----VIKEHANDEGRQLLEQLTEMVTVDAIRM 300
Db 407 AALIWRPRWARLRQPGPVADMLAEGHTMGLVGRGAISTPGRALLDE--AIASADPAIAV 464

QY 301 AGRDMGDFLDEPLEDILSTPEISHVFGDTKLGSVPTPPVLIQVAVHDYLDVSDIDALA 360
Db 465 AA--MTRALPEPIDHFLVQADLTIV-----VPGP-----LQRNL--AKELGTVA 504

QY 361 DSYTAGGANTYHRDLFSEHVSUHLPLSAPMTLRLWLTDRFAGKPLTDHRVTRTWTPTIFNPM 420
Db 505 TVESAGTAMV--YR--ISEQSIIRHALDIGKTRDWMHALFTNHSKTPVPQRLTY--LINDV 558

QY 421 TYA-GMARLAVIAAKV 435
Db 559 ARRHGQLRIGMAASFV 574

RESULT 6
G87178
probable DNA-binding protein [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: G87178
R:Colo, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001

A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: G87178
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-753 <STO>
A:Cross-references: UNIPROT:O9CBEL; GB:AL450380; NID:g13093710; PIDN:CAC31111.1; GSPDB:G
C:Genetics:
C:Gene: ML2156
C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0862c

Query Match 5.0%; Score 116.5; DB 2; Length 753;
Best Local Similarity 25.0%; Pred. No. 1.2;
Matches 124; Conservative 46; Mismatches 209; Indels 117; Gaps 24;

QY 4 PGNLA----GATG---AEWIGRPPHEELQKVRPLLPSPDDPFYFPAGYQHAVPGTVLRS 56
Db 132 PQQVTLIEDISSTGEQIAELIAR--LSQTQRDVLQKLEGSPL----GRTRDAAPGAPPDR 185

QY 57 RDVELAFMGLIPQPVATATQLLYRTNNY-----GNPEAT-----VTTVIIVP-----AELAP 102
Db 186 PVPQLLAMGLLRIRIDADTVILPRRVQVLRGEQSGPTQLTQPVVVSVTTTPNDADAEAG 245

QY 103 GOTCPILLSYQCAIDAMSSRCFPSYALR-----RRAKALG-SLTOMELLMISAALA 151
Db 246 AVIEALHELVDVLETLGSA--PVYELRNGGLGVREFKRLAKATGINEPRGLLLEVTAA 303

QY 152 EGVAVSPDHEGPKLWGSPYEGYRVLDIGIRAAALNSERVGLSPATPIGLWYSGGGLAS 211
Db 304 GLIASGIPDPEPATG--DSPYNAPTATDRFTAMPAPAEKRWHL-----LAS 346

QY 212 ANAAEAC-----GEYAPDDIDVGAVLGSPVGDIGHTFRRLNGTLLAGLPALV-----V 259
Db 347 TWLDLQCRPALIGSRGPDAKSYGALSNSLYSTAAPLDRRLLLGMLAELPAGVGVEAAEAS 406

QY 260 AALQHSYPGLAR-----VIKEHANDEGRQLLEQLTEMVTVDAIRM 300
Db 407 AALIWRPRWARLRQPGPVADMLAEGHTMGLVGRGAISTPGRALLDE--AIASADPAIAV 464

QY 301 AGRDMGDFLDEPLEDILSTPEISHVFGDTKLGSVPTPPVLIQVAVHDYLDVSDIDALA 360
Db 465 AA--MTRALPEPIDHFLVQADLTIV-----VPGP-----LQRNL--AKELGTVA 504

QY 361 DSYTAGGANTYHRDLFSEHVSUHLPLSAPMTLRLWLTDRFAGKPLTDHRVTRTWTPTIFNPM 420
Db 505 TVESAGTAMV--YR--ISEQSIIRHALDIGKTRDWMHALFTNHSKTPVPQRLTY--LINDV 558

QY 421 TYA-GMARLAVIAAKV 435
Db 559 ARRHGQLRIGMAASFV 574

RESULT 7
G82738
gamma-glutamyltranspeptidase XF0984 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: G82738
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-603 <SIM>
A:Cross-references: UNIPROT:Q9PEP4; GB:AE003936; GB:AE003849; NID:g9105908; PIDN:AAF83794
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carter, H.
as-Neto, E.; Docena, C.; El-Dorfy, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
 A:Authors: Martins, E.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
 A:Authors: da Silva, A.C.R.; da Silva, P.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Teuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A:Reference number: A59328
 A:Contents: annotation
 A:Gene: XF0984
 C:Genetics:
 C:Supernatural: gamma-glutamyltransferase

Query Match 4.8%; Score 112.5; DB 2; Length 603;
 Best Local Similarity 22.4%; Pred. No. 1.8;
 Matches 86; Conservative 37; Mismatches 134; Indels 127; Gaps 14;
 QY 47 HAVPGTVLRSRDELAFMG-----LIPOVATATQLLYRTTMYGNPEATVTVIVPAELA 101
 DB 312 HLVVEMRRARRDRFTFFLDGDFVSIQVLTSDYARGLRAAINPEKA-----MPSVLL 366
 QY 102 PQQTCL-----LSYOCADIDAMSSRCFPPSYALRRRAKALGSLQMBELLMISAALAEQNAV- 156
 DB 367 SGHPPTPLEDEETHFSIIIDAENR-----VGATQTVNLLYGSLIPKGTGVL 413
 QY 157 -----SVDPHGPKLWSPYEPGYRVLDTGIRAAALNSRVGLSPATPIGLMGYSGGGLASAAWAEAC 204
 DB 414 LNNEMDDFALQGVNFGVMGYEANAPKPKMLSSMTPTF-----LESADKVAVIGT 467
 QY 205 SGGGLASAAWAEACGEYAPDLIDVCAVLGSPVGDGLHTFRRLNGTLTLAGLPALVVAAL-Q 263
 DB 468 PGGS-----RIITMVL-----LGILCYD-----AGLDAQVAALPR 498
 QY 264 HSYPGLARVIRKEHANDEGRQLLEQIEMTTVDVAVIRMGDMGDFLDPLEDLSTPEIS 323
 DB 499 YHQLWLPDVE-----TEAATPPAEVIKGLTGMHTIKQPSDTIAGGRSS 544
 QY 324 IIVFGDTKLGSVPTPPVLIQVAVHLYLDVSDIDALADSYTAGGANVTYHRDLFSEHVS 383
 DB 545 HVMGNVQTVEMDRRSNVLI-----GGHD-----PRHVVGKAAV----- 577
 QY 384 HPLSAPMTLRLWTDRFACKPLTDH 407
 DB 578 -----WLTDKAVVLPPTDH 591

RESULT 8
 G87681
 penicillin-binding protein AmpH, probable [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
 C:Accession: G87681
 R:Nierman, M.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DoBooy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of Caulobacter crescentus.
 A:Reference number: A87249; MUID:21173698; PMID:11259647
 A:Accession: G87681
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-575 <STO>
 A:Cross-references: UNIPROT:Q9A2R7; GB:AE005673; NID:313425213; PIDN:AAK25451.1; GSPDB:C
 C:Genetics:
 A:Gene: CC3489

Query Match 4.8%; Score 111.5; DB 2; Length 575;
 Best Local Similarity 23.6%; Pred. No. 2;
 Matches 100; Conservative 44; Mismatches 151; Indels 129; Gaps 24;
 QY 30 PLLPSD-----DPFFPPAGYQHAVPGTVLR-----SRDV-----ELAFMGLIPQVPTATOLL 77

DB 153 PRLFTNMAMKPLN-PYADYTEAQLDAFLRDYVLTRDIGGIYEYNSLGV-----GILGLRALA 208
 QY 78 YRTTNMVG-----NPEATVTVIIVPAELAPQOTCPPLIS-YOCAIDAMSSRCFPPSY-- 126
 DB 209 YRAGGDYETVLRQVLTPLGMSDTAIV---LSPAQAARFSSCHNAALEATTTHWDLPSLAG 265
 QY 127 --ALLRRRAKALGSLQMBELLMI-----SAAALAE-----GWAV-SVPD--- 160
 DB 266 AGALURSTADDLLKUTAAELGLVDTPLKAAWADQLVPRPVGGGVEVALGHWVWSTPEGGI 325
 QY 161 --HEGPKGLMGSPYEPGYRVLDTGIRAAALNSRVGLSPATPIGLMGYSGGGLASAAWAEAC 218
 DB 326 VTHSG--GTMGFSQSFVGFNKRKTGLGVVVLSTAGVMGVDDIGLHMTQPLKTPKTRVA 383
 QY 219 GEYAPDLIDVCAVLGSPVGDGLHTFRRLNGTLTLAGLPALVVAALQHSYGLARVIRKEHAN 278
 DB 384 VPLAP-----AAFDKLVGRYAMA-----PCGAVMTIRR--- 410
 QY 279 DEGRQLLEQLTMTT-----VDAVIRMA-----GRDMGDFLDEPIEDILST 319
 DB 411 -DGERMLGQLTGQPTVELFAFSPTTFLKVVDAQITFAVDAEGRGTAVALHQNQN-TTA 468
 QY 320 PEISHVFGDTKLGSVPTPPVLIQVAVHLYLDVSDIDALADSYT-AGGANVTYHR--- 375
 DB 469 PRLAE-----GAEVFAPP---RPAKVAALSVAELDALTGRVALAPGFVTVTRKQS 517
 QY 376 LFSE 379
 DB 518 LFAQ 521

RESULT 9
 G87051
 probable ctpB protein - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C:Accession: G87051
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.R.; Taylor, K.; Whitehead, S.; Bayrell, B.C.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: G87051
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-752 <COL>
 A:Cross-references: UNIPROT:Q10877; GB:Z74410; GB:AL123456; NID:3261600; PIDN:CAA98939.1
 C:Genetics:
 A:Gene: ctpB
 C:Superfamily: Enterococcus copper-transporting ATPase copA; ATPase nucleotide-binding do
 F:21-49/Domain: heavy-metal-associated homology <HMA2>
 F:139-495/Domain: ATPase transduction domain homology <ATT>
 F:569-711/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 4.6%; Score 106.5; DB 2; Length 752;
 Best Local Similarity 25.8%; Pred. No. 6.7;
 Matches 74; Conservative 36; Mismatches 102; Indels 75; Gaps 17;
 QY 43 AGYOHAVP---GTVLSRSDVELAFMGLIPQVPTATOLLYR---TTNMYGNPEATVTV-- 95
 DB 73 AGY-HAAPHTETTVDKRTKD-----PCGAHARRLLRLLVAALFVPLADLSTLPAJ 124
 QY 96 VPAELAPQOTCPPLISYQCAIDAMSSRCFPPSYALRRRAKALGSLQME-LLMISAALAEQW 154
 DB 125 VPSARVFGWGYYLTAAPVTVTAAMPFHSVALR---NARHRTTSMETLISVGIVAATAW 181
 QY 155 AVS-----VPDHGPKLWSPYEPGYRVLDTGIRAAALNSRVGLSPATPIGLMGYS--- 206
 DB 182 SLSSVFGDQPPREG-SGIW-----RAILNSDSIYLEVAAAGTVTVFVLAGRYF 226

QY	207	-----CGLASAWA-----ABACGEVAPDLIDVAGVIGSPVGD-----GHTFRR 245
DB	227	EARAKSKAGSALRALAELGAKNVAVLDP-----GAEIVIPASELKKRQRFVTRPGETI-A 281
QY	246	LGNTLLAGLPALVVAALQ-----HSYPGLARVIKEHANDEGROLLE 286
DB	282	ADGVVVDGSAADMSANTGEAKPVRAIPA-ASVVGTVVMDGLRVIE 327
RESULT 10		
D70634		
probable polyketide synthase cy338 - Mycobacterium tuberculosis (strain H37Rv)		
C:Species: Mycobacterium tuberculosis		
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 12-Jul-2004		
C:Accession: D70634		
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.		
Nature 393, 537-544, 1998		
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.		
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome		
A:Reference number: A70500; MUID:98295987; PMID:9634230		
A:Accession: D70634		
A:Status: preliminary; nucleic acid sequence not shown; translation not shown		
A:Molecule type: DNA		
A:Residues: 1-1402 <COL>		
A:Cross-references: UNIPROT:O86335; GB:284725; GB:AL123456; NID:g3261703; PIDN:CAB06605.		
A:Experimental source: strain H37Rv		
C:Genetics:		
A:Gene: pks6		
C:Keywords: carrier protein		
F:128-520/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OASI>		
F:632-910/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMTI>		
F:1014-1088/Domain: acyl carrier protein homology <ACP1>		
F:1166-1379/Domain: oleoyl-[acyl-carrier-protein] hydrolase homology <ACPH>		
Query Match 4.6%; Score 106.5; DB 2; Length 1402;		
Best Local Similarity 21.1%; Pred. No. 15;		
Matches 120; Conservative 68; Mismatches 188; Indels 193; Gaps 31;		
QY	27	KVRPLLPSP-----DPYFPFAGYQHAVPGTVLRSRD-----VELAPMGLIPOP 70
DB	769	EVAPMLGHVDVIAVNGSPASVVISGANDAVSAIADRLRGQGRVRLHVAHFHSAALMP 828
QY	71	VTATOLLYRTNMGNPEATVTVIIPAEAPG-QTCPLLSY---QCAIDAMSSRCFPSY 126
DB	829	MIA-----EFTA---VAEELSVGLPTIPVISNVTGQVADDFASADYWAR 870
QY	127	ALR-----RRAKGLSLTQMEI-----LMISAALAEWAVSVP-----DHEGP-- 164
DB	871	HIRAVRVFGDSVRSACAGASRFIEVPGGGTSLTEASLADAQIVSVPTLRKDRPEVS 930
QY	165	-----KGL-WGSPYBEGYRV--LDGIRALNSERVGLSPA-----TPGLWG 203
DB	931	VMTAAQAGFVSGMGLDWSVFS--GYRPKRVELPTYAFQHKFWLAPAPSVSDPTAAGQIG 989
QY	204	YSGGG---LASAAAEACEVAPD-----LDIV-----GAVLG--SPVG-DLGHTP--- 243
DB	990	ASDGGELLASSGFAARLAGRSADQLAAAEVVECHAAVILGRGAAGLDAGQAFADSG 1049
QY	244	-----RRINGTLLAGLPALVVAALQHSYP--GLARVIKEHANDEGRQLLEQIEMT 292
DB	1050	FNSLSAVELNRNLTAVTATLPA--TAIFDHPTPTETAQVLITQIDHGSSAAANPAE 1107
QY	293	TVDAVIRM-----ACRDMGDF-----LDEPLEDILTPTPISHVFGDTKL----- 331
DB	1108	RIDALTDLFLQACDAGRDAGKWKVALNSGTRMRSSPVRNNVSKNVALLADGISDVVVI 1167
QY	332	-----GSAPTPPVLIVQAVHDYLDIV- 353
DB	1168	CIPTLVLSDOREYRDIANMTGRHSVYSLTLPGFDSDALPONADMIIVETVSNALIDVY 1227

QY	354	SDIDALADSYTAGCANVTYHRDLSEHVS-----HPLSAPMTLRLWLTDRFPACKPLTDHR 408
DB	1228	GGSCRFVLSGVSOGV-LAY---ALCSHLSVKHQRNPLGLVALIDTYLPSQIANPSMNEG- 1282
QY	409	VRTTPTTFNPM-TYAGMARLAVIAAKVI 436
DB	1283	-----FSPNDTCKGLSREVIRVARM 1303
RESULT 11		
C82767		
phenylalanine-tRNA synthetase beta chain XF0742 [imported] - Xylella fastidiosa (strain 9a)		
C:Species: Xylella fastidiosa		
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004		
C:Accession: C82767		
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequences 406, 151-157, 2000		
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.		
A:Reference number: A82515; MUID:20365717; PMID:10910347		
A:Note: for a complete list of authors see reference number A59328 below		
A:Accession: C82767		
A:Status: preliminary		
A:Molecule type: DNA		
A:Residues: 1-792 <SIM>		
A:Cross-references: UNIPROT:Q9PFD6; GB:AE003916; GB:AE003849; NID:g9105626; PIDN:AAF83552.		
A:Experimental source: strain 9a5c		
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; AJ Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carter, H. as-Neto, E.; Docena, C.; El-Dorcy, H.; Facincani, A.P.; Ferreira, A.J.S.		
submitted to GenBank, June 2000		
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranee, E.E.; Laigre Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E. A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.B.; Zia		
A:Reference number: A59328		
A:Contents: annotation		
C:Genetics:		
A:Gene: XF0742		
C:Superfamily: phenylalanine-tRNA ligase beta chain		
Query Match 4.5%; Score 106; DB 2; Length 792;		
Best Local Similarity 21.0%; Pred. No. 7.8;		
Matches 100; Conservative 69; Mismatches 194; Indels 114; Gaps 25;		
QY	32	LPSDDPFPPFAGYQHAVPGTVLRSRDVELAPMGLIPOPV-----TATQLLYRT 80
DB	136	LPDDAPIGTPLADY-----LALPDASIEIK---LTPNRADCFSVRGIAFDVAAACASEV 186
QY	81	TNMYGNPEATVTVIIPAEAPGQTCPLLSYOCAID---AMSSRCFPSYALRRR---A 132
DB	187	TPLHIDEIPANSAITLPVELHAGANAPRYC--GCVIEGDDPAAPTPTVMARLRSGIRPV 245
QY	133	KALGSLTQMEIIMISAALAEWAVSVPDHEGPKGLWGSYPYEGYRVLDGIRI AALNSERVG 192
DB	246	SLLVDITQVMELEQPM---HAFDVTDLRGPIGVRLSRNDEALKLDGRTVVLDNDPLV 302
QY	193	LSPA--TPIGLMWYSGGGLASAWAAB-----ACGEYAPDLDIVGAVLSPVG---D 238
DB	303	VTDADQPIALAGLIGG---WETRIDTTNVNFLEAAHFAP-VAIMGR--GRKGLGHTD 354
QY	239	LGHTEPRLNGTLLAGI-PALVVAALQHSYPGLARVIKEHANDEGROLLE-QLTMTTVA 296
DB	355	ASHRER-----GVDPALPQAIFAPA---TRILELAGGKPGSLIRVELPEYLPAPA 403
QY	297	VIRMAGRDMGDFLDEPLEDILTPTPISHVFG----DTKLGSAPVTP-----PVLIVQAVHD 348
DB	404	SILLRTRIRALLGIVIDDV-EVERILHALGQVITTOAEGWRVWAPSRFRFDIAIEDLIE 462
QY	349	YLIDVSDIDALADSYTAGGANV-----TTHRDIFS-----EHVSLHPLSAPMTLR 393

Db 463 ELVRIGVYEHLPALPGASHIAMSETRLDWTSVRRQLIARELQETINYAFIDAEILRR 522
QY 304 WLTRDFACKPLTDHVRVTTWPTFNPMYACMARLAVTAAKVI-----TORKLSRR 444
Db 523 W-----QLNTQVMNPLS-----AELAVIRPLLGLVLAALGRNIARQ 562

RESULT 12
JC5290
protein-tyrosine-phosphatase (EC 3.1.3.48) - human
N:Alternate names: Phosphotyrosine phosphatase
C:Species: Homo sapiens (man)
C:Date: 16-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C:Accession: JC5290
R:Wang, B.; Kishihara, K.; Zhang, D.; Hara, H.; Nomoto, K.
Biochem. Biophys. Res. Commun. 231, 77-81, 1997
A:Title: Molecular cloning and characterization of a novel human receptor protein tyrosine kinase.
A:Reference number: JC5290; MUID:97223402; PMID:9070223
A:Accession: JC5290
A:Molecule type: mRNA
A:Residues: 1-1436 <MAN>
A:Cross-references: UNIPROT:O00197; GB:U73727; NID:g1923222; PIDN:AA851343.1; PID:g1923222
C:Comment: This enzyme belongs to type II receptor protein tyrosine phosphatase which mediate and an immunoglobulin-like domains.
C:Genetics:
A:Gene: hPTP-J
C:Superfamily: protein-tyrosine-phosphatase, receptor type mu; fibronectin type III repeat
-tyrosine-phosphatase homology
C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatase
F:22-188/Domain: MAM homology <MAN>
F:203-266/Domain: immunoglobulin homology <IMM>
F:288-366/Domain: fibronectin type III repeat homology <3PR>
F:826-1436/Domain: leukocyte common antigen cytosolic domain homology <LAC>
F:903-1123/Domain: protein-tyrosine-phosphatase homology <PTPI>
F:1193-1418/Domain: protein-tyrosine-phosphatase homology <PTP2>
F:1075/Active site: Cys (phosphotyrosine intermediate) #status predicted
F:1081/Binding site: Cys (phosphotyrosine intermediate) #status predicted
F:1370/Active site: Cys (phosphotyrosine intermediate) #status predicted
F:1376/Binding site: substrate phosphate (Arg) #status predicted

Query Match 4.5%; Score 106; DB 2; Length 1436;
Best Local Similarity 22.6%; Pred. No. 17;
Matches 45; Conservative 27; Mismatches 71; Indels 56; Gaps 6;

QY 85 GNPEATVTTVVPAELAPGQCPLLSYQCAIDAMSSR-----CFPSYALRRRAKA 134
Db 602 GESENTITVLLRPAQ--GRGAPISVQYVIVEERARRLRREPQGCDFPVLTFEALA 658
QY 135 LGSLLTQMELLMSAALAGWAVSPDHGPKLWGSYPGCRVLDGIRAA-----L 186
Db 659 RGLVHYFGAELAAASLPEAMPFTVGDNGTYRGFWNPPLPRKAYLIYFQAASHLKGDTRL 718
QY 187 NSERVLSPATPIGLWYSGGGLASAWAAEACGEYAPDLDT-----VGAVLGSPVGDGL 240
Db 719 NCIRIA-----RKAACKSKRPLEVSQRSEENGILLIGCAGGL- 756

QY 241 HTFRLNLTLLAGLPALVV 259
Db 757 -----AVLLILLGAIIV 768

RESULT 13
TI4591
actinomycin synthetase II acmb [imported] - Streptomyces chrysomallus
C:Species: Streptomyces chrysomallus
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 12-Jul-2004
C:Accession: TI4591
R:Schaeweker, F.; Pfennig, F.; Schroeder, W.; Keller, U.
J. Bacteriol. 180, 2468-2474, 1998
A:Title: Molecular cloning of the actinomycin synthetase gene cluster from Streptomyces
A:Reference number: 218152; MUID:98233744; PMID:9573200

A:Accession: TI4591
A:Status: preliminary; translated from GB/EMBL/DBDJB
A:Molecule type: DNA
A:Residues: 1-2611 <SCH>
A:Cross-references: UNIPROT:O68487; EMBL:AF047717; NID:g3114611; PID:g3114612; PIDN:AA838
C:Genetics:
A:Gene: acmb
C:Superfamily: non-ribosomal peptide synthetase; acetate-CoA ligase homology; acyl carrier
C:Keywords: carrier protein; phosphopantetheine; phosphoprotein
F:509-966/Domain: acetate-CoA ligase homology <ACLI>
F:982-1050/Domain: acyl carrier protein homology <ACPI>
F:1578-2015/Domain: acetate-CoA ligase homology <ACLI2>
F:2031-2098/Domain: acyl carrier protein homology <ACP2>
F:1014,2063/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 4.5%; Score 106; DB 2; Length 2611;
Best Local Similarity 22.5%; Pred. No. 38;
Matches 108; Conservative 52; Mismatches 163; Indels 156; Gaps 26;

QY 9 GATGAEWIGRPPHEELQKV--RELLPSD---DpFYFPPAGYQHAVPGTVLRSDVELA 62
Db 812 GAVGELYVAGP---GLARGYLGRPALTAERFVADP-YGAPGARMYRTGDLVRRNPDSGELE 867
QY 63 FMGLIPOPVTA---TOLLXRTTMMYGNPEATVTTVVPAELAPGQCPLLSYQCAIDAM 118
Db 868 FVGRADHQVKVGRPRIELGEVEAALLAHPDVEQATVIV-REDRPGDT-RLVAVVVGREAL 925
QY 119 SSRCPYSYALLRRRAKALGSITQMELLMSAALAGWAVSPDHGPKLWGSYPGCRV 178
Db 926 -----RPEQVREFTRERL-----PEHMVP----- 944
QY 179 LDGIRALNSERVCLSP-----ATPIGLWYSGGGLASAWAAE--ACGEYAPDLDIVG 229
Db 945 -----AAVVQLERLPLTPNGKLDRAALPEPDFAALAGTGREATPOEQIVC-----DLFT 993
QY 230 AVLGSPPVGDGLGHTFRRLNG-----TLLAGL-----PALVVAALQHS 265
Db 994 QVLGLPRGVDDDFELGCHSLLATRLIAHLRTVGLVGELELSLFEQPTPAVAARLDTA 1053
QY 266 YPG-LARVIEKH-----ANDHGRQL-----LEQITEMTTVDVIRMGDMGDFLDSP-- 312
Db 1054 GPCRALTQRQRPVMPPLSSAQRRRLPFLSTLEGPSATYNIPLVLRLSGR-----LDVPAL 1108
QY 313 ---LEDILTPE-ISHVFGDTKLGSVPTPPVLIVQAVHDYLDIVSDIDALADSYTAGGA 368
Db 1109 CAALGDVVGRHESLRTVPFV---DGTQYQRLTPEAAAPRLTVPPTSEADLPNALKAGA 1165
QY 369 NVTY-----HRDLFS-----EHV-----SLHPLSAPMTLRMLTDRFACK 402
Db 1166 RYAFDLAGEPPLRTELPFELSPREHVLVVAHHAADGMSMEPLS-----RELTFAYAAAR 1219

RESULT 14
A70632
hypothetical protein Rv0435c - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: A70632
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Ganties, S.; Hamlin, N.; Holtroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: A70632
A:Status: preliminary;
A:Molecule type: DNA
A:Residues: 1-728 <COL>
A:Cross-references: UNIPROT:P96281; GB:Z84724; GB:AL123456; NID:g3261708; PIDN:CA806575.1
A:Title: Molecular cloning of the strain H37Rv
C:Genetics:
A:Gene: Rv0435c

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:20:35 ; Search time 55.827 Seconds
(without alignments)
1454.845 Million cell updates/sec

Title: US-10-617-038-9
Perfect score: 1056
Sequence: 1 MIATRDREGATMITFRRLRL.....IRVRNASQWHDLSLFCQTR 210

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1056	100.0	210	8	Adi37288 M. tuberc
2	103.5	9.8	739	8	Ads43317 Bacterial
3	96	9.1	1042	6	Abu33994 Protein e
4	92.5	8.8	256	2	Aar11419 PRV 28K g
5	92.5	8.8	473	2	Aay38686 Neisseria
6	92.5	8.8	473	3	Aay75471 Neisseria
7	90	8.5	361	6	Abu23531 Protein e
8	90	8.5	1088	7	Abm85290 Mouse pro
9	90	8.5	1400	5	Abu05722 M. tuberc
10	90	8.5	5435	4	Aae10145 Streptomy
11	89.5	8.5	238	4	Aau20544 Human sec
12	89.5	8.5	252	5	Abp27393 Streptoco
13	89.5	8.5	393	3	Aay58463 Aureobact
14	89.5	8.5	5588	2	Aay39301 SpnE a po
15	89.5	8.5	5588	4	Aab70969 S. spinos
16	89	8.4	305	6	Abp70430 Amino aci
17	88.5	8.4	1944	8	Adm45915 S. mycaro
18	87.5	8.3	207	7	Adj72231 S. roseop
19	87.5	8.3	895	2	AAR50240 Sequence
20	87.5	8.3	899	8	Abm80426 Tumour-as
21	87.5	8.3	1053	4	Aau28174 Novel hum
22	87.5	8.3	1059	7	Adk65850 Angiogene
23	87.5	8.3	1059	4	Adm88330 Human pro
24	87	8.2	991	4	Abg25255 Novel hum
25	87	8.2	991	6	Abu00874 Polypepti

26	86.5	8.2	456	4	AU48761	Propionib
27	86.5	8.2	456	6	ABM45280	Propionib
28	86.5	8.2	895	2	AAR40929	kDa/15
29	86.5	8.2	895	2	AAR80230	Rabbit 97
30	86.5	8.2	895	2	AAW31588	Rabbit (1
31	86.5	8.2	1047	8	ADR68423	Homologue
32	86.5	8.2	1168	8	ADR68422	Homologue
33	86	8.1	2076	8	ADS73540	Glycopept
34	86	8.1	2468	6	ABU38411	Protein e
35	86	8.1	2468	6	ABP59933	Microbial
36	85.5	8.1	473	3	AAV38687	Neisseria
37	85.5	8.1	473	3	AAV75472	Neisseria
38	85.5	8.1	608	8	ADN25154	Bacterial
39	85	8.0	139	7	ABR83171	ABR83171
40	85	8.0	142	7	ABR83172	DAG-1 int
41	85	8.0	175	7	ABR83179	DAG-1 int
42	85	8.0	410	4	ABB09459	Paraquat
43	85	8.0	478	8	ADN46486	Thermococ
44	85	8.0	895	2	AAR80231	Human dys
45	85	8.0	895	2	AAW31589	Human dys

ALIGNMENTS

RESULT 1
ADI37288
ID ADI37288 standard; protein; 210 AA.
XX
AC ADI37288;
XX
DT 22-APR-2004 (first entry)
XX
DE M. tuberculosis low oxygen induced antigen Rv1733c SEQ ID NO:9.
XX
KW mycobacterial infection; vaccine; tuberculosis;
KW Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;
KW low oxygen induced antigen.
XX
OS Mycobacterium tuberculosis.
XX
PN WO2004006952-A2.
XX
PD 22-JAN-2004.
XX
PF 08-JUL-2003; 2003WO-DK000477.
XX
PR 13-JUL-2002; 2002DK-00001098.
XX
PA (STAT-) STATENS SERUM INST.
XX
PI Andersen P, Rosenkrands I, Stryhn A;
XX
DR WPI: 2004-122778/12.
DR N-PSDB; ADI37333.
XX

Use of one or more polypeptides or their fragments, which are expressed during the latent stage of the mycobacterial infection, and/or nucleic acids encoding the polypeptides, for a therapeutic vaccine against tuberculosis.
Claim 3; SEQ ID NO 9; 76pp; English.
The present invention describes polypeptides or their fragments, which are expressed during the latent stage of a mycobacterial infection, and/or nucleic acids encoding the polypeptides, which are useful for creating a therapeutic vaccine against tuberculosis. Also described: (1) a therapeutic vaccine against tuberculosis comprising one or more polypeptides; (2) a method for treating an animal, including a human being, with tuberculosis caused by virulent mycobacteria, e.g. by Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for immunising an animal, including a human being, against tuberculosis caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis

CC caused by virulent mycobacteria in an animal, including a human being;
CC (5) a method for diagnosing previous or ongoing infection with a virulent
CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
CC infection in a subject. The polypeptides have antibacterial activities,
CC and can be used in vaccines and in gene therapy. The polypeptides are
CC useful for the manufacture of a therapeutic vaccine for treating an
CC individual who is infected by a virulent mycobacterium, e.g. M.
CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
CC The present sequence represents a low oxygen induced antigen, which is
CC used in the exemplification of the present invention.

XX SQ Sequence 210 AA;
Query Match 100.0%; Score 1056; DB 8; Length 210;
Best Local Similarity 100.0%; Pred. No. 2e-101;
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MIATRDREGATMTIFRLPCRTILRVFSRNPLVRGTDRLVAVVLLAVTVSLTTPFA 60
DB 1 MIATRDREGATMTIFRLPCRTILRVFSRNPLVRGTDRLVAVVLLAVTVSLTTPFA 60
QY 61 AAAGTAVQDSRSHVYAHQAQTRHPATATVIDHEGVDSNTTATSAPPRTKITVPARWVN 120
DB 61 AAAGTAVQDSRSHVYAHQAQTRHPATATVIDHEGVDSNTTATSAPPRTKITVPARWVN 120
QY 121 GIERSGEVNAKPGTKSGDRVGIWVDSAGQLVDEPAPPARATADAAALGLWLSVAAVAG 180
DB 121 GIERSGEVNAKPGTKSGDRVGIWVDSAGQLVDEPAPPARATADAAALGLWLSVAAVAG 180
QY 181 ALLALTRAILIRVRNASQWHDIDSLFCTOR 210
DB 181 ALLALTRAILIRVRNASQWHDIDSLFCTOR 210

RESULT 2
ADS43317
ID ADS43317 standard; protein; 739 AA.
XX ADS43317;
AC ADS43317;
DT 02-DEC-2004 (first entry)
XX Bacterial polypeptide #21747.
DE
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX Bacteria.
OS
XX US2003233675-A1.
PN
XX 18-DEC-2003.
PD
XX 20-FEB-2003; 2003US-00369493.
PF
XX 21-FEB-2002; 2002US-0360039P.
PR
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI WPI; 2004-061375/06.
DR
XX New recombinant DNA construct comprising a promoter positioned to provide

PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX Claim 1; SEQ ID NO 21747; 122pp; English.
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 739 AA;

Query Match 9.8%; Score 103.5; DB 8; Length 739;
Best Local Similarity 28.0%; Pred. No. 0.17;
Matches 53; Conservative 17; Mismatches 72; Indels 47; Gaps 8;
QY 21 PCRTILRV-----FSRNPLVRGTDRLVAVVLLAVTVSL-----LTTPFA 60
DB 401 PCALVISVPAMAAALSTGARRGLLVKGGAVIEALARLRLVALDKTGLTPTGRPLVTDLR 460
QY 61 AAAGTAVQDSRSHVYAHQAQTRHPATATVIDH---EGVIDSNTTATSAPPRT----- 109
DB 461 PAAGVAVADLLAHAAAVEAGSSHPALARILARAAEGVTAPATAARVPCVGAEQIAG 520
QY 110 ---KITVPARWVNGIERSGEVNAKPGTKSGDRVGIWVDSAGQ----LVDEPAPPARAT 162
DB 521 ALMRIAAPAA---GNPQAALQAEAKTV-----VELCRDGVFQGLIALRDEPRPEAGA-A 571
QY 163 DAALAAALGL 171
DB 572 LAALRALGL 580

RESULT 3
ABU33994
ID ABU33994 standard; protein; 1042 AA.
XX ABU33994;
AC ABU33994;
DT 19-JUN-2003 (first entry)
XX Protein encoded by Prokaryotic essential gene #19521.
DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
KW Mycobacterium avium.
OS WO200277183-A2.
PN 03-OCT-2002.
PD 21-MAR-2002; 2002WO-US009107.
PF
XX


```

XX WO9924578-A2.
PN
XX
PD
XX
XX
XX 09-OCT-1998; 98WO-IB001665.
XX
XX 06-NOV-1997; 97GB-00023516.
XX 14-NOV-1997; 97GB-00024190.
XX 18-NOV-1997; 97GB-00024386.
XX 27-NOV-1997; 97GB-00025158.
XX 10-DEC-1997; 97GB-00026147.
XX 14-JAN-1998; 98GB-00000759.
XX 01-SEP-1998; 98GB-00019016.
XX (CHIR-) CHIRON SPA.
PA
XX
PI Masignani V, Rappuoli R, Pizza M, Scarlato V, Grandi G;
DR WPI; 1999-327407/27.
DR N-PSDB; AAZ12137.
XX
PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection.
XX
PS Claim 4; Page 237; 524pp; English.
XX
CC Amino acid sequences AAY38499-Y38944 represent Neisseria meningitidis and
CC N. gonorrhoeae antigenic proteins. They are encoded by open reading
CC frames (ORFs) AAZ11972-Z12358. The antigenic proteins, their fragments,
CC their nucleic acids and antibodies are used for diagnosis, prevention (as
CC vaccines) or treatment of Neisseria infections, such as meningitis,
CC septicemia and gonorrhea. Both organisms are closely related. Fragments
CC of the nucleic acids are useful as hybridisation probes and antisense
CC reagents
XX
SQ Sequence 473 AA;
Query Match 8.8%; Score 92.5; DB 2; Length 473;
Best Local Similarity 24.7%; Pred. No. 1.3;
Matches 60; Conservative 31; Mismatches 91; Indels 61; Gaps 12;
QY 10 GATMITFRLR-LPCRTILRV-----FSRNPVVRGTDRLAVVWMLVAVTSLLTIPF 59
Db 113 GIGLVFLFELSFLPIRFLVLRMEGRALAFSAQLV---PKL-AILLPLTVGLLHPA 168
QY 60 AAAAGTAVQ-----DSRSHVYAHQAOTRHPATATVIDHEGVDSNTTATSAP 106
Db 169 NTAVLTVAYALANLAAAAAFLFQNRCLKA----VRHAPFSPAVLHRLRGYGIPIALSSI 224
QY 107 PRKITVPARVVN---GIERSGEVNAKPGTKSGDRV-----GIWVDSAGQLVDEPA 155
Db 225 AYWGLASADRLFLKKYAGLEQLGVYSM--GISFGAALLFQSFSTVMTPTVIFRAIBENA 282
QY 156 PPAR--AIADAALAAALGLWLSVAAGALLAL-----TRAILIRVRNASWQHIDISL 205
Db 283 PPARLSATAESAAALLASALCLTGIFSPLASLLLPENYAARFIVVSCM-----LPPL 335
QY 206 FCT 208
Db 336 FCT 338

```

RESULT 6

AAI75471

ID AAY75471 standard; protein; 473 AA.

XX

AC AAY75471;

XX 21-MAR-2000 (first entry)

XX Neisseria meningitidis ORF 717 protein sequence SEQ ID NO:2416.

XX

```

KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.
XX
OS Neisseria meningitidis.
PN WO9957280-A2.
PD 11-NOV-1999.
XX
XX 30-APR-1999; 99WO-US009346.
XX
XX 01-MAY-1998; 98US-0083758P.
XX 31-JUL-1998; 98US-0094869P.
XX 02-SEP-1998; 98US-0098994P.
XX 02-SEP-1998; 98US-0099062P.
XX 09-OCT-1998; 98US-0103749P.
XX 09-OCT-1998; 98US-0103794P.
XX 09-OCT-1998; 98US-0103796P.
XX 25-FEB-1999; 99US-0121528P.
XX (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mota M;
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarnelli M;
PI Tettelin H, Venter JC;
XX
DR WPI; 2000-062150/05.
DR N-PSDB; AAZ54233.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics.
XX
PS Claim 2; Page 1157; 1453pp; English.
XX
XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
XX represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
XX and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
XX PCR primers used in the exemplification of the present invention. The
XX polypeptides, the polynucleotides, antibodies and compositions of the
XX invention can be used as vaccines, as diagnostic reagents, and as
XX immunogenic compositions. The polypeptides can be used in the manufacture
XX of medicaments for treating or preventing infection due to Neisserial
XX bacteria (e.g. meningitis and septicaemia), to detect the presence of
XX Neisseria bacteria, or to raise antibodies. They may also be used to
XX screen for agonists or antagonists, which may themselves have use as
XX antibacterial agents. The polynucleotides of the invention may also be
XX used in gene therapy protocols
XX
SQ Sequence 473 AA;
Query Match 8.8%; Score 92.5; DB 3; Length 473;
Best Local Similarity 24.7%; Pred. No. 1.3;
Matches 60; Conservative 31; Mismatches 91; Indels 61; Gaps 12;
QY 10 GATMITFRLR-LPCRTILRV-----FSRNPVVRGTDRLAVVWMLVAVTSLLTIPF 59
Db 113 GIGLVFLFELSFLPIRFLVLRMEGRALAFSAQLV---PKL-AILLPLTVGLLHPA 168
QY 60 AAAAGTAVQ-----DSRSHVYAHQAOTRHPATATVIDHEGVDSNTTATSAP 106
Db 169 NTAVLTVAYALANLAAAAAFLFQNRCLKA----VRHAPFSPAVLHRLRGYGIPIALSSI 224
QY 107 PRKITVPARVVN---GIERSGEVNAKPGTKSGDRV-----GIWVDSAGQLVDEPA 155
Db 225 AYWGLASADRLFLKKYAGLEQLGVYSM--GISFGAALLFQSFSTVMTPTVIFRAIBENA 282
QY 156 PPAR--AIADAALAAALGLWLSVAAGALLAL-----TRAILIRVRNASWQHIDISL 205
Db 283 PPARLSATAESAAALLASALCLTGIFSPLASLLLPENYAARFIVVSCM-----LPPL 335
QY 206 FCT 208

```

DB 336 FCT 338

RESULT 7

ABU23531

ID ABU23531 standard; protein; 361 AA.

AC ABU23531;

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #9058.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Bordetella pertussis.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Twawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR N-PSDB; ACA27401.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 51455; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological

CC pathway in which a proliferation-required gene or its gene product lies

CC on a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent

CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antisense nucleic acids are useful for

CC identifying proteins or screening for homologous nucleic acids required

CC for cellular proliferation to isolate candidate molecules for rational

CC drug discovery programs, or for screening homologous nucleic acids

CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of

CC the target prokaryotic essential genes. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 361 AA;

Query Match 8.5%; Score 90; DB 6; Length 361;

Best Local Similarity 24.3%; Pred. No. 1.7;

Matches 54; Conservative 31; Mismatches 81; Indels 56; Gaps 12;

OY 12 TMITRLRLPCRTILRVFSRNPLVRGTDRLKAVVMLAVTVSLLTIPFAAAAGTAVQDSR 71

DB 55 TLAPLHMRL-----TRLLGARP-----RLAAVLTLA-TMLVLLVP-CAMLTASIGDSL 101

OY 72 SHVYAHQAO--TRHPATAT-VID-----HEGVDSNTTATSA--PPRTKITVPAWV 118

DB 102 VNIFHSTGRIVRPDPDFVIELPLIGVRLHEIWLQAMTNLQDVLKPLQPLATAGRWL 161

OY 119 VNGIERSG-----EVNAKPGTKSGDRVGIWVDSAGQ---LVDEPAPPAR 159

DB 162 LAGMASAGLGIAVFLASIAIAGVILVYADPAKRAAHAICGRIAGPEQGVALAQLTAATIR 221

OY 160 AIA-----DAALAALGLWLSVAAVAG--ALLALTRAIL 190

DB 222 AVAQGVIGVALIQAVLAGLGLAGVPAAGLWAIALLICIV 263

RESULT 8

ABM85290

ID ABM85290 standard; protein; 1088 AA.

XX

AC ABM85290;

XX

DT 18-NOV-2004 (first entry)

XX

DE Mouse protein sequence MCP22537.

XX

KW Cytostatic; carcinoma; lymphoma; cancer; murine.

XX

OS Mus musculus.

XX

PN WO2003073826-A2.

XX

PD 12-SEP-2003.

XX

PF 28-FEB-2003; 2003WO-US006235.

XX

PR 01-MAR-2002; 2002US-00087192.

XX

PA (SAGR-) SAGRES DISCOVERY.

XX

PI Morris DW;

XX

DR WPI; 2003-328604/31.

XX

PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma

PT comprises a nucleotide sequence.

XX

PS Claim 5; SEQ ID NO 285; Opp; English.

XX

CC The present invention relates to novel DNA and protein sequences which

CC are associated with carcinomas. The sequences are useful for: (i) for

CC screening drug candidates; (ii) for screening of bioactive agent capable

CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of

CC a bioactive agent capable of modulating the activity of CAP; (iv) for

CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing

CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating

CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;

CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for

CC determining Carcinoma Associated (CA) gene copy number. In addition, the

CC CA genes are useful as DNA vaccines and the CAP are useful as markers of

CC carcinoma including lymphoma. The present sequence is one such CAP. Note:

CC This patent is an equivalent to basic patent US2002182586A1, for which no

CC sequence data was published

FT Domain 9"
 FT 2640..2800
 FT /label= DH9
 FT /note= "Dehydratase domain: part of extender module 9"
 FT 3157..3341
 FT /label= KR9
 FT /note= "Beta-ketoreductase domain: part of extender
 FT module 9"
 FT 3422..3500
 FT /label= AC99
 FT /note= "Acyl carrier protein domain: part of extender
 FT module 9"
 FT 3534..3948
 FT /label= KS10
 FT /note= "Beta-ketosynthase domain: part of extender module
 FT 10"
 FT 4060..4390
 FT /label= AT10
 FT /note= "Acyl transferase domain: part of extender module
 FT 10"
 FT 4413..4597
 FT /label= DH10
 FT /note= "Dehydratase domain: part of extender module 10"
 FT 4900..5078
 FT /label= KR10
 FT /note= "Beta-ketoreductase domain: part of extender
 FT module 10"
 FT 5172..5253
 FT /label= AC910
 FT /note= "Acyl carrier protein domain: part of extender
 FT module 10"
 FT 5302..5555
 FT /label= TE10
 FT /note= "Thiolester domain: part of extender module 10"
 FT XX
 PN WO9946387-A1.
 XX
 PD 16-SEP-1999.
 XX
 PF 16-FEB-1999; 99WO-US003212.
 XX
 PR 09-MAR-1998; 98US-00036987.
 XX
 PA (DOWC) DOW AGROSCIENCES LLC.
 XX
 PI Baltz RH, Broughton MC, Crawford KP, Madduri K, Merlo DJ;
 PI Treadway PU, Turner JR, Waldron C;
 DR WPI; 1999-551414/46.
 DR N-PSDB; AA221501.
 XX
 PT New spinosyn biosynthetic genes from Saccharopolyspora spinosa, useful
 PT for production of insecticidal spinosyn compounds.
 XX
 PS Claim 1; Page 130-148; 190pp; English.
 XX
 CC This is the amino acid sequence of the product of the spnE gene. The
 CC protein is involved in spinosyn biosynthesis. The SpnE gene is one of 23
 CC genes and open reading frames contained in an 80kb DNA sequence AA221501.
 CC Spinosyns are insecticidal microlides which are useful for the control of
 CC arachnids, nematodes and insects. Biosynthesis of spinosyns occurs via
 CC stepwise condensation and modification of carboxylic acid precursors
 CC generating a linear polyketide which is modified further. The DNA
 CC sequence contains a central region of approximately 55kb which has
 CC homology to the DNA encoding the polyketide synthases (PKS) of known
 CC macrolide producers. The spinosyn PKS DNA region consists of 5 ORFs with
 CC stop codons at the end of acyl carrier protein (ACP) domains. Together
 CC the PKS polypeptides (AAV39297-Y39301) of which this sequence is one,
 CC form a complex consisting of an initiator module, spnA, and several
 CC extender modules spnB-spnE. Each extender module adds a specific acetyl
 CC Co-A precursor to a growing polyketide chain, and modifies the beta-keto
 CC group in a specific manner. A module in a PKS polypeptide consists of
 CC several domains with specific functions. The initiator module has an acyl

CC transferase (AT) domain, and an acyl carrier protein (ACP) domain. The
 CC extender modules have the same domains plus a beta-ketosynthase (KS)
 CC domain and optionally a beta-ketoreductase domain, a dehydratase (DH)
 CC domain, and an enoyl reductase (ER) domain. The last extender module
 CC terminates with a thiolester domain. The products of the genes present in
 CC the upstream region the PKS genes have been assigned names spnF-spnS
 CC AAV39302-Y39315 and are responsible for different modifications in
 CC spinosyn biosynthesis. There are also two ORFs ORFL15 and ORFL16 present
 CC immediately upstream of spnS, producing polypeptides AAV39316-Y39317, and
 CC two ORFs ORF1 and ORF2 present downstream of the PKS region producing
 CC polypeptides AAV39318-Y39319. The genes are useful to improve yields of
 CC spinosyns, and for creating new spinosyns e.g. by mutagenesis, or
 CC interruption of steps in spinosyn biosynthesis. The modified spinosyns
 CC may be a new insect control agent or serve as substrates for further
 CC chemical modification and the creation of new semi-synthetic spinosyns.
 CC The genes are also useful to isolate similar sequences from S. spinosa or
 CC other species by hybridization
 XX
 SQ Sequence 5588 AA;
 Query Match 8.5%; Score 89.5; DB 2; Length 5588;
 Best Local Similarity 24.1%; Pred. No. 74;
 Matches 51; Conservative 24; Mismatches 74; Indels 63; Gaps 9;
 Qy 6 RDREGATWITF-----RLRLPCRTI--LRVFSRNPVRGTDRLEAVVM-----LLAVTVS 53
 Db 2920 RPEGGAWLPFSFTGVLSAPAGTISRVRLETRP-----DAISVAVMDSGRLLASIDS 2974
 Qy 54 LLTIPTFAAAAGTAVQDSRSHVYHAQAQTRHPATATVIDHEGVDSNTTATSPRTKITV 113
 Db 2975 L-----RLRSVSSQQLANRDADRDALFEVTWEPVATQSTE 3009
 Qy 114 PARVNVNGIERSGEVN-AKPGTKSGDRVG-----TWVDSAGOLVDEPA 155
 Db 3010 PGRWALLGDTACGKDDLIKLTATDSADRCADLAALAEKLDSSALVPDVVYVYCAGRQAD-PG 3068
 Qy 156 PPARAIADA--ALAALGLWLVAVAAGALLAL 185
 Db 3069 TGAALAEATQOTLALLQNLAEPRLEARLVV 3100
 RESULT 15
 AAB70969
 ID AAB70969 standard; protein; 5588 AA.
 XX
 AC AAB70969;
 XX
 DT 28-AUG-2001 (first entry)
 XX
 DE S. spinosa protein fragment encoded by ORF22, SEQ ID 50.
 XX
 KW Forosamine; trimethylrhannose; polyketide synthase; biosynthesis;
 KW spinosyn; polyketide aglycone; transgenic plant; insect resistance;
 KW macrolide; insecticidal; polyketide synthase.
 XX
 OS Saccharopolyspora spinosa.
 XX
 XX DE19957268-A1.
 XX
 XX 08-MAR-2001.
 XX
 PF 29-NOV-1999; 99DE-01057268.
 XX
 PR 27-AUG-1999; 99DE-01040596.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Eberz G, Moehrlie V, Froede R, Velten R, Salas JA;
 XX
 XX WPI; 2001-267102/28.
 DR N-PSDB; AAF88339.
 XX
 PT New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for

PT recombinant production of insecticidal spinosyns and their derivatives.

PS Claim 57; Page 313-335; 354pp; German.

XX This invention describes a novel method nucleic acid (I) and its encoded
CC polypeptide (II) containing at least one region that encodes an enzymatic
CC activity involved in biosynthesis of spinosyns. (I) are used (i) to
CC identify, inactivate or modulate genes involved in the biosynthesis of
CC (ii); (ii) to generate a library of polyketide synthases; (iii) for
CC adding forosamine or trimethylthamose to a spinosyn or polyketide
CC aglycone; and (iv) for recombinant production of the corresponding
CC enzymes, which are used for production of (II), their precursors or
CC derivatives, including production of transgenic plants that express (II)
CC and thus have increased resistance to insects. (I) are also useful as
CC markers for sequencing of the Saccharopolyspora spinosa genome. (II) are
CC macrolides with insecticidal, but not antibacterial, activity, and can
CC also be used to raise specific antibodies, useful for identifying
CC expression clones in a gene bank. Cells transformed with (I) may produce
CC (II) at significantly increased levels or produce new derivatives of
CC (II). This sequence represents an S. spinosa polyketide synthase

XX SQ Sequence 5588 AA;

Query Match 8.5%; Score 89.5; DB 4; Length 5588;

Best Local Similarity 24.1%; Pred. No. 74; Mismatches 74; Indels 63; Gaps 9;

Matches 51; Conservative 24; Mismatches 74; Indels 63; Gaps 9;

QY 6 RDREGATMITF-----RLRLPCTI--LRVFSRNPLVRGTDRLAEVVM-----LLAVTVS 53

Db 2920 RPEQGGAWLPFSFTGTVELSAPACTISRVRLETRP-----DAISVAVMDESGRLASIDS 2974

QY 54 LITIPFAAAGTAVQDSRSHVYAHQOTRHPATATVIDHEGVDSNTTATSAPPRKTIIV 113

Db 2975 L-----RLRSVSSGQLANRDVARDALFEVTEPVATQSTE 3009

QY 114 PARWVYNGIERSEVN-AKPGTKSGDRVG-----IWVDSAGQLVDEPA 155

Db 3010 PGRWALLGDTACKDGLIILATDSADRCADLAALAEKLDSSALVPDVPVVCAGEQAD-PG 3068

QY 156 PPARAIADA--ALAAALGLWLSVAAVAGALLAL 185

Db 3069 TGAALAEYQQITLALLQAWLAEPRLAEARLVV 3100

Search completed: October 15, 2005, 04:26:26

Job time : 57.827 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 01:42:44 ; Search time 17.2419 Seconds
(without alignments)
909.197 Million cell updates/sec

Title: US-10-617-038-9
Perfect score: 1056
Sequence: 1 MIATRDREGATMITFRLRL.....IRVRNASQWHDIDSLFCTQR 210

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/prodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCITUS COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	89.5	8.5	5588	3 US-09-036-987A-6	Sequence 6, Appli
2	89.5	8.5	5588	3 US-09-370-700-6	Sequence 6, Appli
3	89.5	8.5	5588	3 US-09-603-207-6	Sequence 6, Appli
4	85	8.0	895	1 US-08-123-161A-8	Sequence 8, Appli
5	85	8.0	895	1 US-08-483-278-8	Sequence 8, Appli
6	85	8.0	895	4 US-09-949-016-6490	Sequence 6490, Ap
7	85	8.0	895	4 US-09-949-016-7178	Sequence 7178, Ap
8	85	8.0	2736	4 US-09-252-991A-30227	Sequence 30227, A
9	84.5	8.0	222	4 US-09-902-540-10599	Sequence 10599, A
10	84.5	8.0	1996	2 US-08-804-227C-9	Sequence 9, Appli
11	84.5	8.0	1996	2 US-08-804-198-3	Sequence 3, Appli
12	83	7.9	281	4 US-09-902-540-13631	Sequence 13631, A
13	82.5	7.8	425	4 US-09-910-430-32	Sequence 32, Appl
14	81	7.7	367	4 US-09-252-991A-23330	Sequence 23330, A
15	80	7.6	437	4 US-09-252-991A-25331	Sequence 25331, A
16	80	7.6	502	4 US-09-724-797-90	Sequence 90, Appl
17	80	7.6	610	4 US-09-252-991A-20096	Sequence 20096, A
18	80	7.6	1399	4 US-09-388-221B-4	Sequence 4, Appli
19	80	7.6	1424	4 US-09-388-221B-12	Sequence 12, Appl
20	80	7.6	1443	4 US-09-388-221B-6	Sequence 6, Appli
21	80	7.6	1454	4 US-09-388-221B-10	Sequence 10, Appl
22	80	7.6	1473	4 US-09-388-221B-2	Sequence 2, Appli
23	79.5	7.5	218	4 US-09-252-991A-25202	Sequence 25202, A
24	79.5	7.5	412	4 US-09-079-030-124	Sequence 124, App
25	79.5	7.5	470	4 US-09-902-540-12670	Sequence 12670, A
26	79	7.5	349	4 US-09-252-991A-21699	Sequence 21699, A
27	79	7.5	903	4 US-09-902-540-11595	Sequence 11595, A

28 78.5 7.4 119 3 US-09-036-574-2 Sequence 2, Appli
29 78.5 7.4 119 4 US-08-454-294A-2 Sequence 2, Appli
30 78.5 7.4 263 4 US-09-252-991A-26300 Sequence 26300, A
31 78.5 7.4 263 4 US-09-252-991A-28078 Sequence 28078, A
32 78 7.4 217 4 US-09-902-540-12408 Sequence 12408, A
33 78 7.4 284 4 US-09-252-991A-30616 Sequence 30616, A
34 78 7.4 590 4 US-09-902-540-9740 Sequence 9740, Ap
35 78 7.4 1749 4 US-09-640-419C-28 Sequence 28, Appl
36 77.5 7.3 209 4 US-09-388-221B-19 Sequence 19, Appl
37 77.5 7.3 410 4 US-09-252-991A-21004 Sequence 21004, A
38 77.5 7.3 501 4 US-09-902-540-11195 Sequence 11195, A
39 77 7.3 267 4 US-09-902-540-13611 Sequence 13611, A
40 77 7.3 390 4 US-09-252-991A-20123 Sequence 20123, A
41 77 7.3 618 4 US-09-252-991A-27666 Sequence 27666, A
42 76.5 7.2 136 4 US-09-513-999C-6261 Sequence 6261, Ap
43 76.5 7.2 401 4 US-09-489-039A-8285 Sequence 8285, Ap
44 76 7.2 272 4 US-09-248-796A-16689 Sequence 16689, A
45 76 7.2 317 4 US-09-134-000C-5669 Sequence 5669, Ap

ALIGNMENTS

RESULT 1
US-09-036-987A-6
; Sequence 6, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clave
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; NUMBER OF INVENTION: Production
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow AgroSciences LLC Patent Department
; STREET: 9310 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,987A
FILING DATE: 09-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 5588 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-036-987A-6

Query Match 8.5%; Score 89.5; DB 3; Length 5588;
Best Local Similarity 24.1%; Pred. No. 11;
Matches 51; Conservative 24; Mismatches 74; Indels 63; Gaps 9;

QY 6 RDREGATMITF-----RLRLPCRTI--LRVFSRNPLVRGTDRLAEAVM-----LLAVTVS 53
Db 2920 RPEQGGAWLPFSFTGVLSAPAGTISRVRLETRRP-----DAISVAVMDESGRLLASIDS 2974
QY 54 LRTIPFAAAGTAVQDSRSRVYHQAQTRHPATATVIDHEGVDSNTTATSPAPRTKITV 113
Db 2975 L-----RLRSVSSGQLANRDALFEVTPWEPVATQSTE 3009
QY 114 PARVVNGIERSGEVNV-AKPGTKSGDRVG-----IWDVSAGQLVDEPA 155
Db 3010 PGRWALLGDTACGKDDLIKATDSADRCADLAALAEKLDSSALVDPVVVYCAGEQAD-PG 3068
QY 156 PPARAIADA--ALAAALGLWLSVAAGALLAL 185
Db 3069 TGAALAEQTQTLALLQAWLAEPRLAEARLVV 3100

RESULT 2

US-09-370-700-6
; Sequence 6, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/370,700
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/36987
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 5588
; TYPE: PRT
; ORGANISM: Saccharopolyspora spinosa
US-09-370-700-6

Query Match 8.5%; Score 89.5; DB 3; Length 5588;
Best Local Similarity 24.1%; Pred. No. 11;
Matches 51; Conservative 24; Mismatches 74; Indels 63; Gaps 9;
QY 6 RDREGATMITF-----RLRLPCRTI--LRVFSRNPLVRGTDRLAEAVM-----LLAVTVS 53
Db 2920 RPEQGGAWLPFSFTGVLSAPAGTISRVRLETRRP-----DAISVAVMDESGRLLASIDS 2974
QY 54 LRTIPFAAAGTAVQDSRSRVYHQAQTRHPATATVIDHEGVDSNTTATSPAPRTKITV 113
Db 2975 L-----RLRSVSSGQLANRDALFEVTPWEPVATQSTE 3009
QY 114 PARVVNGIERSGEVNV-AKPGTKSGDRVG-----IWDVSAGQLVDEPA 155
Db 3010 PGRWALLGDTACGKDDLIKATDSADRCADLAALAEKLDSSALVDPVVVYCAGEQAD-PG 3068
QY 156 PPARAIADA--ALAAALGLWLSVAAGALLAL 185
Db 3069 TGAALAEQTQTLALLQAWLAEPRLAEARLVV 3100

RESULT 3

US-09-603-207-6
; Sequence 6, Application US/09603207B
; Patent No. 6521406
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P

; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/603,207B
; CURRENT FILING DATE: 2000-06-23
; EARLIER APPLICATION NUMBER: 09/370,700
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 5588
; TYPE: PRT
; ORGANISM: Saccharopolyspora spinosa
US-09-603-207-6

Query Match 8.5%; Score 89.5; DB 4; Length 5588;
Best Local Similarity 24.1%; Pred. No. 11;
Matches 51; Conservative 24; Mismatches 74; Indels 63; Gaps 9;
QY 6 RDREGATMITF-----RLRLPCRTI--LRVFSRNPLVRGTDRLAEAVM-----LLAVTVS 53
Db 2920 RPEQGGAWLPFSFTGVLSAPAGTISRVRLETRRP-----DAISVAVMDESGRLLASIDS 2974
QY 54 LRTIPFAAAGTAVQDSRSRVYHQAQTRHPATATVIDHEGVDSNTTATSPAPRTKITV 113
Db 2975 L-----RLRSVSSGQLANRDALFEVTPWEPVATQSTE 3009
QY 114 PARVVNGIERSGEVNV-AKPGTKSGDRVG-----IWDVSAGQLVDEPA 155
Db 3010 PGRWALLGDTACGKDDLIKATDSADRCADLAALAEKLDSSALVDPVVVYCAGEQAD-PG 3068
QY 156 PPARAIADA--ALAAALGLWLSVAAGALLAL 185
Db 3069 TGAALAEQTQTLALLQAWLAEPRLAEARLVV 3100

RESULT 4

US-08-123-161A-8
; Sequence 8, Application US/08123161A
; Patent No. 5449616
; GENERAL INFORMATION:
; APPLICANT: Campbell, Kevin P.
; APPLICANT: Roberts, Steven L.
; APPLICANT: Anderson, Richard D.
; APPLICANT: Ibraghimov, Oxana B.
; APPLICANT: Yang, Bin
; TITLE OF INVENTION: NUCLEIC ACID ENCODING DYSTROPHIN-ASSOCIATED
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kevin M. Farrell, P.C.
; STREET: P.O. Box 999
; CITY: York Harbor
; STATE: ME
; COUNTRY: USA
; ZIP: 03911
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/123,161A
; FILING DATE: 16-SEP-93
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/946,234
; FILING DATE: 14-SEP-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Farrell, Kevin M.

; REGISTRATION NUMBER: 35,505
; REFERENCE/DOCKET NUMBER: UIR989-11A4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (207) 363-0558
; TELEFAX: (207) 363-0528
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 895 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-123-161A-8

Query Match 8.0%; Score 85; DB 1; Length 895;
Best Local Similarity 31.1%; Pred. No. 2.4;
Matches 28; Conservative 6; Mismatches 26; Indels 30; Gaps 4;

QY 78 QAOHRPATATVIDHEGVDSNTTATSAP---PRTKITVP-----A 115
DB 428 RVSTPKPATPST-----DSTTTTTRRPTKKPRTPRVVTTKVSITRLETASPPTRI 480
QY 116 RRVVNGIERSGEVNAKPGTKSG-DRVGIWV 144
DB 481 RTTSGVPRGGEPNORPELKNHIDRVDAWV 510

RESULT 5

US-08-483-278-8
; Sequence 8, Application US/08483278
; Patent No. 5686073

; GENERAL INFORMATION:
; APPLICANT: Campbell, Kevin P.
; APPLICANT: Ibraghimov, Oxana B.
; APPLICANT: Ervasti, James M.
; APPLICANT: Leveille, Cynthia J.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING DYSTROPHIN-ASSOCIATED
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESS: Kevin M. Farrell, P.C.
; STREET: P.O. Box 999
; CITY: York Harbor
; STATE: ME
; COUNTRY: USA
; ZIP: 03911

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,278
; FILING DATE:
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/123,161
; FILING DATE: 16-SEP-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Farrell, Kevin M.
; REGISTRATION NUMBER: 35,505
; REFERENCE/DOCKET NUMBER: UIR989-11A5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (207) 363-0558
; TELEFAX: (207) 363-0528
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 895 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

Query Match 8.0%; Score 85; DB 1; Length 895;

Best Local Similarity 31.1%; Pred. No. 2.4;
Matches 28; Conservative 6; Mismatches 26; Indels 30; Gaps 4;
QY 78 QAOHRPATATVIDHEGVDSNTTATSAP---PRTKITVP-----A 115
DB 428 RVSTPKPATPST-----DSTTTTTRRPTKKPRTPRVVTTKVSITRLETASPPTRI 480
QY 116 RRVVNGIERSGEVNAKPGTKSG-DRVGIWV 144
DB 481 RTTSGVPRGGEPNORPELKNHIDRVDAWV 510

RESULT 6

US-09-949-016-6490
; Sequence 6490, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6490
; LENGTH: 895
; TYPE: PRT
; ORGANISM: Human

US-09-949-016-6490

Query Match 8.0%; Score 85; DB 4; Length 895;
Best Local Similarity 31.1%; Pred. No. 2.4;
Matches 28; Conservative 6; Mismatches 26; Indels 30; Gaps 4;

QY 78 QAOHRPATATVIDHEGVDSNTTATSAP---PRTKITVP-----A 115
DB 428 RVSTPKPATPST-----DSTTTTTRRPTKKPRTPRVVTTKVSITRLETASPPTRI 480
QY 116 RRVVNGIERSGEVNAKPGTKSG-DRVGIWV 144
DB 481 RTTSGVPRGGEPNORPELKNHIDRVDAWV 510

RESULT 7

US-09-949-016-7178
; Sequence 7178, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7178
; LENGTH: 920
; TYPE: PRT
; ORGANISM: Human

US-09-949-016-7178

Query Match 8.0%; Score 85; DB 4; Length 920;
Best Local Similarity 31.1%; Pred. No. 2.5;
Matches 28; Conservative 6; Mismatches 26; Indels 30; Gaps 4;
Qy 78 QAQTRHPATATVIDHEGVDSNTTATSAP---PRTKITVP-----A 115
Db 453 RVSTPKPATPST-----DSTTTTTRPTKKPRTPRPVPTTKVSTIRLETASPPTRI 505
Qy 116 RWVNGIERSGEVNAKPKTSG-DRVGIIWV 144
Db 506 RTTSGVRGGEPNQRPKLNHIDRVDAWV 535

RESULT 8

US-09-252-991A-30227
; Sequence 30227, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30227
; LENGTH: 2736
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30227

Query Match 8.0%; Score 85; DB 4; Length 2736;
Best Local Similarity 26.9%; Pred. No. 12;
Matches 49; Conservative 21; Mismatches 84; Indels 28; Gaps 6;
Qy 5 TRDREGATMITRRLPCTRTILRVFSRNPLVGRDRLLEAVVMLLAVTVSLTIPFAAAG 64
Db 1466 TADGSGNMSFTPGIPLDGTVMVVARSP--SNVDSAPAVI-----TVDGVAPAA 1513
Qy 65 TAVODSRSHVVAHQATHPATATVIDHEGVDSNTTATSAPPRTKITVPARVVNGIER 124
Db 1514 PVIDPSNGTEISGTAB-----AGATVILTDGGNPIGQATADGSGNWTFTPTPLANGT-- 1567
Qy 125 SGEVNAKPKTSGDRVGIWVDSAGOLVDEPAPPARAIADAALGLWLVAVAAGALLA 184
Db 1568 --VINAQAQDPAAGTSG---PASVTVDIAIAPPAPVINPSN---GVVISGTAEAGATVI 1617
Qy 185 LT 186
Db 1618 LT 1619

RESULT 9

US-09-902-540-10599
; Sequence 10599, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10599
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10599

Query Match 8.0%; Score 84.5; DB 4; Length 222;
Best Local Similarity 29.2%; Pred. No. 0.35;
Matches 50; Conservative 20; Mismatches 64; Indels 37; Gaps 10;
Qy 20 LPCTRIILRVFSRNPL--VRGTDRLLEAVVML-----LAVTVSLTIPFAAAGTAVQDSRS 72
Db 4 LPARFILRLVDRTVAGRVLET-DLSYSVREARGVAVVPSTATVPDPASPDLTIVDERV 62
Qy 73 HVIYAHQAQTRHPATATVIDHEGVDSNTTATSAP---PRTKITVPARVVNGIERSGEVNA 130
Db 63 YLLA-----RGPDTLLILD---VLNAAAGTAPDAGTTESVRI-VSALFVPAGASELEVIPI 113
Qy 131 KPG-----TKSGDR-VGIWVDSAGOLV-----DEPAPPARAIAD 163
Db 114 GPRGNLVAVTSGDEAVAIFDEVGOLVAQVQVGDNDPNQSPFFGLAAD 164

RESULT 10

US-08-804-227C-9
; Sequence 9, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: DeHoff, Bradley S.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1996 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-804-227C-9

Query Match 8.0%; Score 84.5; DB 2; Length 1996;
Best Local Similarity 27.8%; Pred. No. 8.9;
Matches 47; Conservative 23; Mismatches 70; Indels 29; Gaps 9;
Qy 37 GTDRLEAVVMLLAVTVSLTIPFAAAGTAVQDSRSRVYAHQAQTRHPATATVIDHEGVI 96
Db 63 GTDAVSAPPVDRGWDVEGLYDPODEAVGRS-----YVREGGLHSAEFDAEFFGI- 113


```
QY 97 DSNNTATSPRTKITVPARWVNGVNGIERSGEVNAK-PGTSKGRVCIWVDSAGQLVDEPA 155
Db 114 -SPREAAAMPQORLLLETSW--EALERAGIVPASLRGTRTGVTGVMYDDYGSRFDS-A 169
QY 156 PP-----ARATADAALA-ALGL---WLSV-AAVAGALLALTRAI 189
Db 170 PPEYEGVNGSAGSIASGRVAALGLEGPALTVDTCSSSVLHVALV 218

RESULT 11
US-08-804-198-3
; Sequence 3, Application US/08804198
; Patent No. 5945320
; GENERAL INFORMATION:
; APPLICANT: Burgett, Stanley G.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rao, Nagaraja R.
; APPLICANT: Richardson, Mark A.
; APPLICANT: Rostock, Paul R., Jr.
; TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: PAUL R. CANTRELL, 1138
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,198
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CANTRELL, PAUL R.
; REGISTRATION NUMBER: 36,470
; REFERENCE/DOCKET NUMBER: P9113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3885
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1996 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-804-198-3

Query Match 8.0%; Score 84.5; DB 2; Length 1996;
Best Local Similarity 27.8%; Pred. No. 8.9;
Matches 47; Conservative 23; Mismatches 70; Indels 29; Gaps 9;

QY 37 GTDRLEAVVLLAVTVSLTIPFAAAGTAVQDSRSHVVAHQATRHPATATVIDHEGV 96
Db 63 GTDAVSAPFVDRGWDVEGLVDPDPEAVGRS-----YVREGGLHLSAAEFDAEPFGI- 113
QY 97 DSNNTATSPRTKITVPARWVNGVNGIERSGEVNAK-PGTSKGRVCIWVDSAGQLVDEPA 155
Db 114 -SPREAAAMPQORLLLETSW--EALERAGIVPASLRGTRTGVTGVMYDDYGSRFDS-A 169
QY 156 PP-----ARATADAALA-ALGL---WLSV-AAVAGALLALTRAI 189
Db 170 PPEYEGVNGSAGSIASGRVAALGLEGPALTVDTCSSSVLHVALV 218

RESULT 12
US-09-902-540-13631
; Sequence 13631, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
```

```
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13631
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13631

Query Match 7.9%; Score 83; DB 4; Length 281;
Best Local Similarity 26.0%; Pred. No. 0.73;
Matches 54; Conservative 23; Mismatches 63; Indels 68; Gaps 10;

QY 45 VMLLAVTVSLTIPFAAAGTAVQDSRSHVVAHQATRHPATATVIDHEGVDSNNTAT 104
Db 29 TAAALALSCMTSGTEAAESASKRESRAEAKAKARIIEVAALVDQHD-----KKEAAA 83
QY 105 APPRTKITVPARWVNGVNGIERSGEVNAK-----PQTKSGDRV----- 140
Db 84 ATRRAK-EVEAQ-----TIORLQEIADQVARRKQAAAAQAGEGTPSGPEVQPEPTAVVPSL 138
QY 141 -----GIWVDSAGQLVD--BPAP-PA-----RAIADAAALALGLWLSVA 176
Db 139 SGAGDESPESTGAEADVAGATPDQAKPSPIPADTPDARLEQIRIDVEAAAG----- 192
QY 177 AVAGALLALTRAILIRVRNASHQHDIDS 204
Db 193 --ADAGTSLAKAITALVMEAT--KDLDS 216

RESULT 13
US-09-910-430-32
; Sequence 32, Application US/09910430
; Patent No. 6794166
; GENERAL INFORMATION:
; APPLICANT: Godfroi, Edmond
; APPLICANT: Bollen, Alex
; APPLICANT: Lebouille, Gerard
; TITLE OF INVENTION: IDENTIFICATION AND MOLECULAR CHARACTERIZATION OF
; TITLE OF INVENTION: PROTEINS, EXPRESSED IN THE IXODES RICINUS SALIVARY
; TITLE OF INVENTION: GLANDS
; FILE REFERENCE: VANM229.001CP1
; CURRENT APPLICATION NUMBER: US/09/910,430
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: PCT/BE00/00061
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: GB 9913425.6
; PRIOR FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Ixodes ricinus
US-09-910-430-32

Query Match 7.8%; Score 82.5; DB 4; Length 425;
Best Local Similarity 34.0%; Pred. No. 1.5;
Matches 32; Conservative 9; Mismatches 36; Indels 17; Gaps 3;

QY 126 GEVNAKPGTSKGRVGIW---VDSAGQLVDEP-----APPARAIADAAAL 169
Db 168 GAQGSFPRARCATAGLCCFVVDSDNQLVLQFCMEVLQVGVKAGKAGPGSPVLOAAALLAW 227
QY 170 GLWLSVAVAGALLALTRAILIRVRNASHQHDID 203
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:14:10 ; Search time 50.2192 Seconds
(without alignments)
1742.576 Million cell updates/sec

Title: US-10-617-038-9
Perfect score: 1056
Sequence: 1 MIATTRDREGATMTFRLRL.....IRVRNASWQHDIDSLFCTOR 210

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1056	100.0	210	15	US-10-617-038-9
2	169	16.0	209	14	US-10-156-761-9051
3	167	15.8	201	14	US-10-156-761-8969
4	149.5	14.2	288	14	US-10-156-761-8840
5	103.5	9.8	739	15	US-10-369-493-21747
6	96	9.1	1042	15	US-10-282-122A-61918
7	94	8.7	3723	17	US-10-844-716-6
8	91.5	8.7	772	16	US-10-425-115-225624
9	90.5	8.6	466	14	US-10-156-761-8526
10	90	8.5	361	15	US-10-282-122A-51455
11	90	8.5	1088	13	US-10-087-192-285
					Sequence 9, Appli
					Sequence 9051, Ap
					Sequence 8969, Ap
					Sequence 8840, Ap
					Sequence 21747, A
					Sequence 61918, A
					Sequence 6, Appli
					Sequence 225624,
					Sequence 8526, Ap
					Sequence 51455, A
					Sequence 285, App

ALIGNMENTS

RESULT 1

US-10-617-038-9
; Sequence 9, Application US/10617038
; Publication No. US20040057963A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Peter
; APPLICANT: Rosenkrands, Ida
; APPLICANT: Stryhn, Anette
; TITLE OF INVENTION: Therapeutic TB Vaccine
; FILE REFERENCE: SSISAUSA
; CURRENT APPLICATION NUMBER: US/10/617,038
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: DK PA 2002 01098
; PRIOR FILING DATE: 2002-07-13
; PRIOR APPLICATION NUMBER: US 60/401,725
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-617-038-9

Query Match 100.0%; Score 1056; DB 15; Length 210;
Best Local Similarity 100.0%; Pred. No. 8.1e-100;
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MIATTRDREGATMTFRLRLPCTILRVFSRNPILVRGTDRLEAVMLLVTSLLTIPFA 60
Db 1 MIATTRDREGATMTFRLRLPCTILRVFSRNPILVRGTDRLEAVMLLVTSLLTIPFA 60
QY 61 AAAGTAVQSRSHVYAHQATRHPTATVIDHEGVTDSTTATSAPPRKITVPAWVNV 120

Db 61 AAAGTAVQDSRSHVYAHQAQTRHPATATVIDHEGVDSNTTATSAPPRKTIIVPARWVN 120
QY 121 GIERSEVNAKPGTSGDRVGIWDSAGOLVDEPAPPARAIADAALGLWLSVAAYAG 180
Db 121 GIERSEVNAKPGTSGDRVGIWDSAGOLVDEPAPPARAIADAALGLWLSVAAYAG 180
QY 181 ALLALTRAILIRVRNASQWHDIDSIFCTOR 210
Db 181 ALLALTRAILIRVRNASQWHDIDSIFCTOR 210

RESULT 2

US-10-156-761-9051
; Sequence 9051, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9051
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9051

Query Match 16.0%; Score 169; DB 14; Length 209;
Best Local Similarity 32.2%; Pred. No. 7.9e-09;
Matches 56; Conservative 19; Mismatches 87; Indels 12; Gaps 6;
QY 31 RNPLVGTDRLEAVVMLLAVTVSLTIPFAAAGTAVQDSRSHVYAHQAQTRHPATATVI 90
Db 32 RNQLRRSDVLEAWLFSATVTVLC--GVLITGLAAHSVEQGLAREAEWRPVRLVA 88
QY 91 DHEGVDSNTTATSAPPRKTIIVPARW-VVNGIERSEVNAKPGTKSGDRVGIWDSAGQ 149
Db 89 --EDAPESAAASGADR--VMKVRWTTANGSTHSGQARVAAGSATGTPVTVTWTRDGL 144
QY 150 LVDEPAPPARAIADAALGLWLSVAAYAGALLA--LTRAILIRVRNASQWHD 201
Db 145 LVTKVPVPE--AQVRGVTVGLLVCSAAAGPFVGGRLVRGLRRRMRWDED 196

RESULT 3

US-10-156-761-8969
; Sequence 8969, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8969
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8969

Query Match 15.8%; Score 167; DB 14; Length 201;
Best Local Similarity 30.6%; Pred. No. 1.2e-08;
Matches 55; Conservative 23; Mismatches 90; Indels 12; Gaps 5;
QY 32 NPLVGTDRLEAVVMLLAVTVSLTIP--FAAAGTAVQDSRSHVYAHQAQTRHPATATVI 90
Db 13 NPLRRTTDLIEAWVTLTALLILVAVPVIGAVGVAQDALQQSVRDQRRARHETVATVV 72
QY 91 D--HEGVDSNTTATSAPPRKTIIVPARWVVNGIERSEVNAKPGT-KSGDRVGIWDS 147
Db 73 KKLNRGGLDPPDPTSSARDAHSRVLAAWTGPDGSAAHGAVLADLKTPHRGDHFLLMTDQ 132
QY 148 GOLVDEPAPPARAIADAALGLWLSVAAYAGALLALTRAI----LIRVRNASQWHDID 203
Db 133 GRMVGRLDTATATTHAMLAGFG----AAAMSAGLVEGGRLLIVRMVRRVRYARWDOAWD 188

RESULT 4

US-10-156-761-8840
; Sequence 8840, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8840
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8840

Query Match 14.2%; Score 149.5; DB 14; Length 288;
Best Local Similarity 29.0%; Pred. No. 1.2e-06;
Matches 54; Conservative 26; Mismatches 75; Indels 31; Gaps 9;
QY 31 RNPLVGTDRLEAVVMLLAVTVSLTIPFAAAA--GTAVQDSRSHVYAHQAQTRHPATAT 88
Db 106 RNPLRRSDVVEAWTALVAVLFFVAAPLAGAAVGMWAYDSSHSLATAQRAERKH-VSAT 164
QY 89 VIDHEGVDSNTTATSAPPRK-----ITVPARWVNGI-ERSEVNAKPGTKSGDRVG 141
Db 165 LVEN-----APAAVPSVQDSKEPLYQVKVWTFEGRGARTGPAWVPAGSLRGERTD 215
QY 142 IWDSAGOLVDEPAPPARAIADAAL-----AALGLWLSVAAYAGALLA--LTRAILIRVRN 195
Db 216 VWLDGRGRSV---APP---VTDLVWQHAVTSGVMTTGGVVLVLLARIAERRIARHRM 269
QY 196 ASWQHD 201
Db 270 AEWEDE 275

```
RESULT 5
US-10-369-493-21747
; Sequence 21747, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 21747
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Rhodobacter capsulatus
US-10-369-493-21747

Query Match          9.8%; Score 103.5; DB 15; Length 739;
Best Local Similarity 28.0%; Pred. No. 0.22;
Matches 53; Conservative 17; Mismatches 72; Indels 47; Gaps 8;

QY 21 PCRTILRV-----FSRNPLVRGTDRLAEVVMLLAVTVSL-----LTIPFA 60
Db 401 PCALVISVPMAAALSTGARRLLVKGGAVERALRLRLVALDKTGTTPGRPLVTDLR 460
QY 61 AAAGTAVQDSRSVHYAHOAQTRHPATATVIDH---EGVIDSNTTATSPART----- 109
Db 461 PRAQVAVADLLAHAAAVEAGSSHPLARAILARAAGVTPAPATAARVIPGVGAEAQIAG 520
QY 110 ---KITVPARVVVNGIERGEVNAKPGTKSGDRVGIWVDSAGO-----LVDEPAPPARATA 162
Db 521 ALWRIAAPAA---GNPQAAALQAEAKTV-----VELCRDGVPGQGLIALRLDEPRPEAGA-A 571

QY 163 DAALALGL 171
Db 572 LAALRALGL 580

RESULT 6
US-10-282-122A-61918
; Sequence 61918, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITPA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26

US-10-844-716-6
; Sequence 6, Application US/10844716
; Publication No. US20050003409A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Chengjin
; APPLICANT: Chaleff, Deborah T.
; APPLICANT: Ruppen, Mark E.
; APPLICANT: Stephens, Jerome
; TITLE OF INVENTION: Cloning Genes From Streptomyces Cyaneogriseus Subsp.
; TITLE OF INVENTION: Noncyanogenus for Biosynthesis of Antibiotics and Methods of Use
; FILE REFERENCE: AM100484
; CURRENT APPLICATION NUMBER: US/10/844,716
; CURRENT FILING DATE: 2004-05-13
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 3723
; TYPE: PRT
; ORGANISM: bacteria
US-10-844-716-6

Query Match          8.9%; Score 94; DB 17; Length 3723;
Best Local Similarity 28.4%; Pred. No. 18;
Matches 54; Conservative 20; Mismatches 78; Indels 38; Gaps 8;

QY 23 RTILRVFSRNPLVRGTDRLAEVVMLLAVTVSLLTIPFAAAAGTAVQDSRSVHYA--HQAQ 80
Db 3313 RHLLLVSRRGMAAAGAEKLCBELGQAGVSVS-----AGCDVAD-RAQVAALLEQVP 3363
QY 81 TRHPATATVIDHEGVDSNTTATSPARTKITVPARVVVNGIERGEVNAKPGTKSGDRV 140
Db 3364 ABHPLTA-VVHTAGVLDDAT--VTCLDRNKIDAVILGAKVDGALHLHELTA-----GMDL 3414
```

```
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61918
; LENGTH: 1042
; TYPE: PRT
; ORGANISM: Mycobacterium avium
US-10-282-122A-61918

Query Match          9.1%; Score 96; DB 15; Length 1042;
Best Local Similarity 21.8%; Pred. No. 2.1;
Matches 34; Conservative 24; Mismatches 48; Indels 50; Gaps 5;

QY 78 OQTRHPATATVIDHEGVDSNTTATSPARTKITVPARW-----VVNGIERSGEVNAKP 132
Db 634 EALTRTP-TAVGDDGATVSARTSPVAAPSRPNTAPHTYSKKHPLITDMVRNTTLNGPK 692
QY 133 GTK-----SGDRVGIWVDSAGQVDEPAPPARATAAALALGLWLSV 175
Db 693 STKDVRLHVLNLPEDAVSYEAGDALGWPGRNSDELVDE-----WLAV 734
QY 176 AAVAG-----ALLALTRAILIRVRNASQWHD 202
Db 735 TGLDGTQTPVEVGEHGLMSLSALTERTIEIAHISRD 770

RESULT 7
US-10-844-716-6
; Sequence 6, Application US/10844716
; Publication No. US20050003409A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Chengjin
; APPLICANT: Chaleff, Deborah T.
; APPLICANT: Ruppen, Mark E.
; APPLICANT: Stephens, Jerome
; TITLE OF INVENTION: Cloning Genes From Streptomyces Cyaneogriseus Subsp.
; TITLE OF INVENTION: Noncyanogenus for Biosynthesis of Antibiotics and Methods of Use
; FILE REFERENCE: AM100484
; CURRENT APPLICATION NUMBER: US/10/844,716
; CURRENT FILING DATE: 2004-05-13
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 3723
; TYPE: PRT
; ORGANISM: bacteria
US-10-844-716-6

Query Match          8.9%; Score 94; DB 17; Length 3723;
Best Local Similarity 28.4%; Pred. No. 18;
Matches 54; Conservative 20; Mismatches 78; Indels 38; Gaps 8;

QY 23 RTILRVFSRNPLVRGTDRLAEVVMLLAVTVSLLTIPFAAAAGTAVQDSRSVHYA--HQAQ 80
Db 3313 RHLLLVSRRGMAAAGAEKLCBELGQAGVSVS-----AGCDVAD-RAQVAALLEQVP 3363
QY 81 TRHPATATVIDHEGVDSNTTATSPARTKITVPARVVVNGIERGEVNAKPGTKSGDRV 140
Db 3364 ABHPLTA-VVHTAGVLDDAT--VTCLDRNKIDAVILGAKVDGALHLHELTA-----GMDL 3414
```


US-10-282-122A-51455

Query Match 8.5%; Score 90; DB 15; Length 361;
Best Local Similarity 24.3%; Pred. No. 2.1;
Matches 54; Conservative 31; Mismatches 81; Indels 56; Gaps 12;
QY 12 TMTFRLRLPCRTILRVFNSPLVRGTDRLAEVWMLLATVTSLLTIPFAAAAGTAVQDSR 71
DB 55 TLAPLHMRL-----TRLLGARP-----RLAAVLTLA-TMLVLLVP-GAMLTASIGDSL 101
QY 72 SHVYAHQAO--TRHPATA--VID-----HEGVDSNTTATSA--PPRTKITVPAWV 118
DB 102 VNIHFHSTGRIVRIPAPPPFVIELPLGLVRLHEIWMQNTNLQDVLKPLQPHLATAGWL 161
QY 119 VNGIERSG-----EYNKPGTKSGDRVGIWVDSAGQ-----LVDEPAPPAR 159
DB 162 LAGMASAGLGIADVFLASIAIGVILVYADPAKRAAHAIGVRIAGPEQVALAQLTAATIR 221
QY 160 ATA-----DAAALGLWLSVAAG--ALLATRAIL 190
DB 222 AVAQGVIGVALIOAVLAGLGLLAVPAAGLWALVALLICIV 263

RESULT 11

US-10-087-192-285
; Sequence 285, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 285
; LENGTH: 1088
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-087-192-285

Query Match 8.5%; Score 90; DB 13; Length 1088;
Best Local Similarity 28.3%; Pred. No. 9.1;
Matches 32; Conservative 9; Mismatches 42; Indels 30; Gaps 4;
QY 55 LTIPFAAAAGTAVQDSRSHVYAHQAOTRHPATATVIDHEGVDSNTTATSAP---PRTKI 111
DB 429 LIIPGVETAVITPTTTTKKPRVSTPAPTST-----DSSITTRTKKPRTR 481
QY 112 TYP-----ARVWNGIERSGEVNAKPGTKSG-DRVGIWV 144
DB 482 PVPRVTTKAPITRLETASPPTIRTTTSGVPRGGEPNQRPCLKNHIDRDVAWV 534

RESULT 12

US-10-080-170-373
; Sequence 373, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123

; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 373
; LENGTH: 1400
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-373

Query Match 8.5%; Score 90; DB 14; Length 1400;
Best Local Similarity 23.7%; Pred. No. 13;
Matches 67; Conservative 24; Mismatches 74; Indels 118; Gaps 13;
QY 28 VFSRNPPLVRGTDRLB-----AVVMLL-----AVTVSLLTTPFAAAAGT 65
DB 1057 VIDRNAL--GFDQLKPPGLAEVVLVSAGGAPIAPADAARNRALTVDCHDGPVVAAGR 1114
QY 66 AVQDS-RSHVYA-----HQAOTRHPATATVID-----HEGVID 97
DB 1115 FVHTSIRITVIGALLDGPVVAALPCREPIALPAGQOELLISPCAAAFVVDGAQLSTPGAGL 1174
QY 98 SNTTATSA-----PPRTKITVP-----ARWV-----VNGIE 123
DB 1175 SSATVTSAETGAWGPTHREVRVPESATSRVLVVPESINSWGVAARTSTGARLTPIAVNGWQ 1234
QY 124 RSGEVNA-KPGT-----KSGDRVGI-----WVDSAGOLVDEPAPPARA 160
DB 1235 QAWVVPAGNPGTITLTTFAPNSLYRASLAIGLALLPLALLAFWRTGRRQLADRPPTPPWRP 1294
QY 161 TADAALAALGLWLSVAAGALL-----ALTRAILRVR 194
DB 1295 GAWAAAGVLAAGAVIASIAGVVMGTALGVRYALRRRLDR 1337

RESULT 13

US-10-080-170-373
; Sequence 373, Application US/10080170
; Publication No. US20040121322A9
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 373
; LENGTH: 1400
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-373

Query Match 8.5%; Score 90; DB 16; Length 1400;
Best Local Similarity 23.7%; Pred. No. 13;
Matches 67; Conservative 24; Mismatches 74; Indels 118; Gaps 13;
QY 28 VFSRNPPLVRGTDRLB-----AVVMLL-----AVTVSLLTTPFAAAAGT 65
DB 1057 VIDRNAL--GFDQLKPPGLAEVVLVSAGGAPIAPADAARNRALTVDCHDGPVVAAGR 1114
QY 66 AVQDS-RSHVYA-----HQAOTRHPATATVID-----HEGVID 97
DB 1115 FVHTSIRITVIGALLDGPVVAALPCREPIALPAGQOELLISPCAAAFVVDGAQLSTPGAGL 1174
QY 98 SNTTATSA-----PPRTKITVP-----ARWV-----VNGIE 123
DB 1175 SSATVTSAETGAWGPTHREVRVPESATSRVLVVPESINSWGVAARTSTGARLTPIAVNGWQ 1234
QY 124 RSGEVNA-KPGT-----KSGDRVGI-----WVDSAGOLVDEPAPPARA 160

```
Db 1235 QAWVVPAGNPGTITLTFAPNSLYRASLAIGLALLPLLAFWRTGRRQLADRPPTPPWRP 1294
Qy 161 IADAAALGLWLWSVAAGALL-----ALTRAILIRV 194
Db 1295 GAWAAAGVLAAGAVIAGVVMGTALGVRYALRRRLRDR 1337

RESULT 14
US-10-468-356-373
; Sequence 373, Application US/10468356
; Publication No. US20040197896A1
; GENERAL INFORMATION:
; APPLICANT: COLE, STEWART
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 05394.0019
; CURRENT APPLICATION NUMBER: US/10/468,356
; CURRENT FILING DATE: 2003-08-19
; PRIOR FILING DATE: 10/080,170
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 655
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 373
; LENGTH: 1400
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-468-356-373

Query Match 8.5%; Score 90; DB 16; Length 1400;
Best Local Similarity 23.7%; Pred. No. 13;
Matches 67; Conservative 24; Mismatches 74; Indels 118; Gaps 13;

Qy 28 VFSRNPLVRGTDRL-----AVVMLL-----AVTSLLTIPFAAAAGT 65
Db 1057 VIDRNAL--GFDQLKPPGLAEVWVLSAGGAPIAPADAARNRALTVDCHGPPVAVAGR 1114
Qy 66 AVQDS-RSHVYA-----HQAQTRHPATATVID-----HEGVID 97
Db 1115 FVHTSIRTVGALLDGEFVAALPCEREPIALPAGQOELLISPGAAFFVVDGAQLSTPGAGL 1174
Qy 98 SNTTATSA-----PPRTKITVP-----ARWV-----VNGIE 123
Db 1175 SSATVTSATGANGTHREVVPESATSRVLVPESINSGWVARTSTGARLTPIAVNGWQ 1234
Qy 124 RSGEVNA-KPGT-----KSGDRVGI-----WVDSAGQLVDEPAPPARA 160
Db 1235 QAWVVPAGNPGTITLTFAPNSLYRASLAIGLALLPLLAFWRTGRRQLADRPPTPPWRP 1294
Qy 161 IADAAALGLWLWSVAAGALL-----ALTRAILIRV 194
Db 1295 GAWAAAGVLAAGAVIAGVVMGTALGVRYALRRRLRDR 1337
```

```
RESULT 15
US-10-203-295-38
; Sequence 38, Application US/10203295
; Publication No. US20040115762A1
; GENERAL INFORMATION:
; APPLICANT: Zotchev, Sergey Borisovich
; APPLICANT: Sekurova, Olga Nikolayivna
; APPLICANT: Fjaervik, Epsen
; APPLICANT: Brautaset, Trygve
; APPLICANT: Strom, Arne Reidar
; APPLICANT: Valla, Svein
; APPLICANT: Ellingsen, Trond Erling
; APPLICANT: Slette, Havard
; APPLICANT: Gulliksen, Ole-Martin
; TITLE OF INVENTION: Novel genes encoding a nystatin polyketide synthase and their
; TITLE OF INVENTION: manipulation and utility
```

```
; FILE REFERENCE: 1181-265
; CURRENT APPLICATION NUMBER: US/10/203,295
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: PCT/GB 01/00509
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: GB 0002840.7
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: GB 0008786.6
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: GB 0009387.2
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 38
; LENGTH: 5435
; TYPE: PRT
; ORGANISM: Streptomyces noursei ATCC 11455
US-10-203-295-38

Query Match 8.5%; Score 90; DB 16; Length 5435;
Best Local Similarity 26.0%; Pred. No. 78;
Matches 63; Conservative 24; Mismatches 79; Indels 76; Gaps 14;

Qy 16 FR-LRLPCRTILRVFSRNPLVRGTDRLLEAVVMLLAVTVSLL-TIPFAAAAGTAVQDSRSH 73
Db 1093 FRGLRAVWRRDTEIYAEVALPEGTADADAFGLHPALFDVHLSTLFPASADG----DQSL 1148
Qy 74 V-YAHQAQTRHPATA-----TVIDHEG-----VIDSNTTATSAP----- 106
Db 1149 LPFAWNGVSLHAAGADALRVITSCGPDAVEITAVDPQGRPVSVESLTLRAAGPDAGTA 1208
Qy 107 -----PRT--KITVPAEWMVNGIERSGEVN-----AKPCTKSGDRVGIWVD 145
Db 1209 DHRADAGSLFRMDWTPRTVTHAPATPATWAVLGTDPIGLTEALTAAAGPDTVTGLKDG--VD 1266
Qy 146 SAGQLV---DEPAP-----PARAIAD-----AALAALGLWLWSVAAGA-CALIAL 185
Db 1267 ALGELTAGDDRPPVDVAVPLRGATDHGPAGAHDLTITVLLALLQEWLAEEFARSRLLV 1326
Qy 186 TR 187
Db 1327 TR 1328

Search completed: October 15, 2005, 04:15:00
Job time : 51.2192 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:59:30 ; Search time 11.7178 Seconds
(without alignments)
1724.342 Million cell updates/sec

Title: US-10-617-038-9
Perfect score: 1056
Sequence: 1 MIATTRDREGATMITFRLRL.....IRVRNASWQHIDSLSFCTQR 210

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1056	100.0	210	2 H70687	hypothetical prote
2	204.5	19.4	193	2 T36959	hypothetical prote
3	153.5	14.5	456	2 T35474	50kD proline rich
4	105.5	10.0	459	2 T29443	probable two-compo
5	92.5	8.8	256	1 WMBEPN	28K protein - suid
6	92.5	8.8	473	2 C81039	lipopolysaccharide
7	90.5	8.6	815	2 C84287	hypothetical prote
8	90	8.5	297	2 H87523	hypothetical prote
9	90	8.5	650	2 S59630	dystroglycan alpha
10	90	8.5	1400	2 J70963	hypothetical prote
11	89.5	8.5	586	2 JC4825	glucan 1,4-beta-gl
12	89	8.4	274	2 B84361	hypothetical prote
13	89	8.4	436	2 T34980	probable two-compo
14	87	8.2	257	2 B84346	hypothetical prote
15	86.5	8.2	895	2 S20582	dystrophin-associa
16	86.5	8.2	1168	2 E70726	probable polyketid
17	86	8.1	2468	2 A83412	hypothetical prote
18	85.5	8.1	473	2 C81984	probable lipopolys
19	85	8.0	410	2 JC7715	paragut resistant
20	85	8.0	509	2 C70624	probable two compo
21	85	8.0	895	2 I54343	dystroglycan - hum
22	85	8.0	1098	2 G70697	probable arabinosy
23	84.5	8.0	1350	2 AF2005	RNA polymerase bec
24	84	8.0	278	2 D75585	urease accessory p
25	84	8.0	485	2 T35628	probable penicilli
26	83.5	7.9	483	2 H72640	hypothetical prote
27	83.5	7.9	884	2 C70729	hypothetical prote
28	83.5	7.9	1616	2 G70668	polyketide synthas
29	82.5	7.8	606	2 T37217	probable ATP/GTP b

ALIGNMENTS

RESULT 1

H70687
hypothetical protein Rv1733c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: H70687
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A: Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A: Reference number: A70500; MUID: 98295987; PMID: 96344230
A: Accession: H70687
A: Status: Preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-210 <COL>
A: Cross-references: UNIPROT: P71991; GB: Z81360; GB: AL123456; NID: g3261654; PIDN: CAB03696.1
A: Experimental source: strain H37RV
C: Genetix:
A: Gene: Rv1733c
C: Superfamily: Mycobacterium tuberculosis hypothetical protein Rv1733c

Query Match	100.0%	Score 1056;	DB 2;	Length 210;
Best Local Similarity	100.0%	Pred. No. 2.3e-82;		
Matches 210;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MIATTRDREGATMITFRLRLPCRTILRVFSRNPLVRGTDRLVAVVLLAVTVSLLTIPFA	60	
DB	1	MIATTRDREGATMITFRLRLPCRTILRVFSRNPLVRGTDRLVAVVLLAVTVSLLTIPFA	60	
QY	61	AAAGTAVODSRSHVYAHQAOTRHPATATVIDHEGVDSNTTATTSAPPRTKITVPARVVN	120	
DB	61	AAAGTAVODSRSHVYAHQAOTRHPATATVIDHEGVDSNTTATTSAPPRTKITVPARVVN	120	
QY	121	GIERSEVNAKPGTKSGDRVGIWVDSAGQLVDEPAPPARAIADAALAAALGLMLSVAAVAG	180	
DB	121	GIERSEVNAKPGTKSGDRVGIWVDSAGQLVDEPAPPARAIADAALAAALGLMLSVAAVAG	180	
QY	181	ALLALTRAILIRVRNASWQHIDSLSFCTQR	210	
DB	181	ALLALTRAILIRVRNASWQHIDSLSFCTQR	210	

RESULT 2

T36959
hypothetical protein SCJ1.26 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36959
R: Seeger, K.J.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1999


```

RESULT 8
H87523
hypothetical protein CC2217 [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C/Accession: H87523
R/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heide-
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.B.; Haft, D.;
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fra-
J. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: H87523
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-297 <STO>
A/Cross-references: UNIPROT:Q9A677; GB:AF005673; NID:G13423722; PIDN:AAK24188.1
C/Genetics:
A/Gene: CC2217

```

```

Query Match      8.5%; Score 90; DB 2; Length 297;
Best Local Similarity 26.6%; Pred.No. 2.7;
Matches 55; Conservative 20; Mismatches 70; Indels 62; Gaps 9;

QY    33 PLYRGTD-RLKAVVMLLAVTVSLTTI-PFAAAGTAVQDSRSHVYAHQAQTRHPATATVI 90
       ||| : | | : | : ||| : | | | : | | | : | | | : | | | :
Db    14 PLLPPRDARDGALVFVVAVLCFLACLTAFAALAANRA-----AHGWTAQLSGSATVV 65

QY    91 DHESVIDSNTTATSPAPRTKITVPARVVNGVIERSCGVNAKPCTKSDCRVGIVWDSAGQL 150
       ||| : | | : | : ||| : | | | : | | | : | | | :
Db    66 VR---ARSNETPDSAAARAETA-----GVRGVTEAQALPREKAELLAPWI-GAEAL 115

QY    151 VDE-----PAPPARAIADAAL-----AAL 169
       ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    116 VDPLPTPLVTLEDPKAPTAGILDRALRAAGVDATVDDHSRWIADIERAANIARPAAL 175

QY    170 GLWLSSVAAVAGALLAL--TRAILIRVRN 195
       ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    176 GVFTALIAATAAVIAFATRAGLAARRD 202

RESULT 9
SY59630
dystroglycan alpha precursor - mouse (fragment)
N/Alternate names: laminin-binding glycoprotein
C/Species: Mus musculus (house mouse)
C/Date: 15-Feb-1996 #sequence revision 31-Dec-2000 #text_change 09-Jul-2004
C/Accession: S59630; I48336; S47180; S52849
R/Branchaccio, A.; Rueegg, M.A.; Engel, J.
```

Matrix Biol. 14, 681-685, 1995
A:Title: Cloning and sequencing of mouse skeletal muscle alpha-dystroglycan.
A:Reference number: S59630; MUID:97210754; PMID:9057818
A:Accession: S59630
A:Molecule type: mRNA
A:Residues: 1-650 <BRA>
A:Cross-references: UNIPROT:Q62165; EMBL:X86073; NID:G2924286; PIDN:CAA60031.1; PID:G2924286
R:Gorecki, D.C.; Derry, J.M.; Barnard, E.A.
Hum. Mol. Genet. 3, 1589-1597, 1994
A:Title: Dystroglycan: brain localisation and chromosome mapping in the mouse.
A:Reference number: I48336; MUID:95135414; PMID:7833916
A:Accession: I48336
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 352-447, 'S', 449, 'E', 451-598, 'PH', 601-650 <GOR>
A:Cross-references: EMBL:Z34532; NID:G506465; PIDN:CAA84293.1; PID:G506466
C:Genetics:
A:Gene: DAG1
C:Keywords: glycoprotein
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-650/Product: dystroglycan alpha (fragment) #status predicted <MAT>
F:647/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 8.5%; Score 90; DB 2; Length 650;
Best Local Similarity 28.3%; Pred. No. 6.5;
Matches 32; Conservative 9; Mismatches 42; Indels 30; Gaps 4;

Qy 55 LTIFPAAAGTAVQDSRHVHAQAQTRHPATATVIDHEGVDSNTTATSA-----PRTKI 111
Db 403 LTIFGVVEPTAVITPTTTTKKRVSTPKPATST-----DSSTTTTRPTKKRPTPR 455

Qy 112 TVP-----ARWVNGIERSGEVNAKPKTSG-DRVGTWV 144
Db 456 PVPRVTTKAPTRLETASPPTRITTSVGRGSGPNQRPKLKHIDRVDAWV 508

RESULT 10
B70963
hypothetical protein Rv0236C - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: B70963
R:Colle, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Cole, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: B70963
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1400 <COL>
A:Cross-references: UNIPROT:P96419; GB:Z92669; GB:AL123456; NID:G3242271; PIDN:CAB07017.
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv0236C
C:Superfamily: Mycobacterium leprae probable integral membrane protein

Query Match 8.5%; Score 90; DB 2; Length 1400;
Best Local Similarity 23.7%; Pred. No. 15;
Matches 67; Conservative 24; Mismatches 74; Indels 118; Gaps 13;

Qy 28 VFSRNPLVRGTDRLLE-----AVVMLL-----AVTVSLLTIPFAAAGT 65
Db 1057 VIDNAL--GFDQLKPGLEAVVVLVSAGAPIADAARNRALTVDCHGPPVAVAGR 1114

Qy 66 AVQDS-RSHVVA-----HQQTRHPATATVID-----HEGVID 97
Db 1115 FVHTSIRTVGALLDGEFVAALPCEREPALPAGQQLLISFGAAFFVVDGQLSTPAGL 1174

Qy 98 SNTTATSA-----PPRTKITVP-----ARWV-----VNGIE 123

Db 1175 SSATVTSAGTGWGPTHREVRVPESATSRVLVVPESINSGWVARTSTGARLTPIAVNGWQ 1234
Qy 124 RSEVNA-KGQT-----KSGDRVGI-----WVDSAGQLVDEPAPPARA 160
Db 1235 QAWVVPAGNPGTITLTFTAPNSLYRASIAIGLALLPLLALLAFWRTGRRQLADRPPTPMRP 1294

Qy 161 IADALAALGLWLSAAVAGALL-----ALTRAILIRVR 194
Db 1295 GAWAAAGVLAAGAVIASIAGVWVGTAAGVRYALRRERLRDR 1337

RESULT 11
JC4825
glucan 1,4-beta-glucosidase (EC 3.2.1.74) - Microbispora bispora
N:Alternate names: exo-1,4-beta-glucosidase
C:Species: Microbispora bispora
C:Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 09-Jul-2004
C:Accession: JC4825
R:Goyal, A.K.; Eveleigh, D.E.
Gene 172, 93-98, 1996
A:Title: Cloning, sequencing and analysis of the ggh-A gene encoding a 1,4-beta-D-glucan
A:Reference number: JC4825; MUID:96257260; PMID:8654999
A:Accession: JC4825
A:Molecule type: DNA
A:Residues: 1-986 <GOY>
A:Cross-references: UNIPROT:Q59506; GB:L061134; NID:G149825; PIDN:AA836835.1; PID:G149826
C:Comment: This enzyme belongs to group 3 cellulose hydrolyzing enzymes.
C:Genetics:
A:Gene: ggh-A
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
F:890-983/Domain: Clostridium xylanase A repeat homology <XA>

Query Match 8.5%; Score 89.5; DB 2; Length 986;
Best Local Similarity 24.4%; Pred. No. 11;
Matches 53; Conservative 25; Mismatches 84; Indels 55; Gaps 10;

Qy 30 SRNPLVRGTDRLLEAVVMLLA-----VTVSLLTIPFAAAAGTAVQ-DSRSHVVAH 77
Db 143 AQPVVWGLNMAPVNVLLRDPGRNREGYSEDPLLTGAIAIAYGSGIQGDDPDHLRA- 201

Qy 78 QAQTRHPATATVIDHEGVDSNTTATSA-----ITVPARVWVNGIERSGE-----V 128
Db 202 -----APTLLKHLANNNEIRDTTSSNLPVRKHEYVEAPFRAAITAGAAATGVTAYNLV 256

Qy 129 NAKPGTKS---GDRVGTWVD-SAGQLVDEPAP-----PARATADALAALGL--- 171
Db 257 NGRPATVNPDLNDTVRTWTDRLNVTDAGAPNNLVGSAQYFATLADADAALAKAGIDSF 316

Qy 172 ----WLSVAAVAGALLALTRAILIRVRNAGWQHDIDS 204
Db 317 TTDETNSAPTITAKTALSQGLLT-----EQDIDT 346

RESULT 12
B84361
hypothetical protein Vng2113C [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: B84361
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Liu
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: B84361
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-274 <STO>
A:Cross-references: UNIPROT:Q9HNG6; GB:AE004437; NID:G10581527; PIDN:ANG20254.1; GSPDH:GH
C:Genetics:
A:Gene: VNG2113C

Query Match 8.4%; Score 89; DB 2; Length 274;
Best Local Similarity 21.5%; Pred. No. 3;
Matches 38; Conservative 41; Mismatches 80; Indels 18; Gaps 7;
QY 34 LVRCGRLEAVVLLAVTVSLTIFPAAAGTAVQDSR---SHVYAHQOHRPA-TATV 89
DB 92 IVIGTAARTTISGLIAGLVVWARPFEIGWIDALADAEGTVDDITVTVTRVNPSEEVV 151
QY 90 IDHEGVDSNTTATAPPRTKITVPARVWVNGIERSGEVNAKPGTKSGDRVGIWVDSAGQ 149
DB 152 IPENVLNATINRTSLDRLSLVAV-----GVDYDADIET-----AEDVIEDAITDYSR 201
QY 150 LVDEPAPPA--RAIDADAALA-ALGWLVSVAAGALLALTRAILIRVRNASWQHDDID 203
DB 202 ILDTAPQVVPVPTALGDSAIVLSCREWFTHIPSPAKRSMA-TAAVVRVAVKTDADID 257
RESULT 13
T34980
probable two-component sensor - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C:Accession: T34980
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1998
A:Reference number: Z21550
A:Accession: T34980
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-436 <OLI>
A:Cross-references: UNIPROT:086661; EMBL:AL031182; PIDN:CAA20157.1; GSPDB:GN00070; SCOPD:1.1.1.1
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOPDB:SC4A2.05
C:Superfamily: hypothetical protein H11707; sensor histidine kinase homology
Query Match 8.4%; Score 89; DB 2; Length 436;
Best Local Similarity 28.6%; Pred. No. 5.1;
Matches 50; Conservative 20; Mismatches 75; Indels 30; Gaps 9;
QY 40 RLEAVVLLA-VTVSLTIFPAAAGTAVQDSRSHVYAHQAT--RHPATATVIDHEGVI 96
DB 4 RVVRVALTAALVAVILLAVPLALAI-----RSSLYADQDRTLEAALSAARVRSPPDT 56
QY 97 DSNNTTATAPP--RTKITVPAWVWVG-----IERSGEVNAKPGTKSGDR-VGIW 143
DB 57 TGDVPELPAPPADGRGLVDPSLRIRAGTGHGTAVRRAISAEVVRGPGDLVAVP 116
QY 144 VDSAGQLVD--EPAPPARAIAADAAALGLWLSVAAGALLALTRAILIRVRNA 196
DB 117 VSHAERVIGVWRASSPAAAVDRVLIA---WAVLAACVCA--LALTAVLVARRQA 166
RESULT 14
B84346
hypothetical protein Vng1952h [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: B84346
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laaky, S.; Leithausner, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo, Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: B84346
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-257 <STO>
A:Cross-references: UNIPROT:Q9HNT6; GB:AE004437; NID:g10581389; PIDN:AAG20134.1; GSPDB:C
C:Genetics:

A:Gene: VNG1952H
Query Match 8.2%; Score 87; DB 2; Length 257;
Best Local Similarity 28.0%; Pred. No. 4.1;
Matches 42; Conservative 10; Mismatches 60; Indels 38; Gaps 6;
QY 28 VFSRNPLVRCGRLEAVVLLAVTVSLTIFPAAAGTAVQDSRSHVYAHQ-----AQY 81
DB 1 MFDSMDTTRTIRRTAVVTVAVIAAL-----AAAGSTAAASPPPTPHRFPGTVSDAD 56
QY 82 RHPA---TATVIDHEGVDSNTTATAPPRTKITVPA-----RWV 118
DB 57 GHPAGNVTVVHNGSVASNTTADSG--YYDLTVPADAVSAGRSVTVTARNQSAQFEWT 114
QY 119 VNGIERSGEVNAKPGTKSGDRVGIWVDSAG 148
DB 115 AAG---STEVEFTLGPSSGSDVEDGTDSSG 141
RESULT 15
S20582
dystrophin-associated glycoprotein precursor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S20582
R:Braghimov-Beskrovnaya, O.; Ervasti, J.M.; Leveille, C.J.; Slaughter, C.A.; Sernett, S.
Nature 355, 696-702, 1992
A:Title: Primary structure of dystrophin-associated glycoproteins linking dystrophin to t
A:Reference number: S20582; MUID:92158075; PMID:1741056
A:Accession: S20582
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-895 <IBR>
A:Cross-references: UNIPROT:Q28685; EMBL:X64393; NID:g1545; PIDN:CAA45732.1; PID:g1546
Query Match 8.2%; Score 86.5; DB 2; Length 895;
Best Local Similarity 28.7%; Pred. No. 18;
Matches 47; Conservative 11; Mismatches 61; Indels 45; Gaps 9;
QY 4 TTRDREGATMITFRRLRCPRTILRVFSRNPVLRGTDRLVAVVLLAVTVSLTIFPAAAA 63
DB 369 TTRTR-GAIIQTPTLG-PIQP-TRVSDAGTVVSGQIR-----ATVTIPGYVEPTAVAT 418
QY 64 GTAVQDSRSHVYAHQOQTRHPATATVIDHEGVDSNTTATSAP---PRKITVTP----- 114
DB 419 PPTTTTKKPRV-----STPKPATPST-----DSSATTTTRRPTKKRTPRPVPRVTTKA 466
QY 115 -----ARWVNGIERSGEVNAKPGTKSG-DRVGIWV 144
DB 467 PITRLETASPPTRIRTTTSGVPRGGEPNORPELKNHIDRVDAWV 510
Search completed: October 15, 2005, 04:38:27
Job time : 12.7178 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:24:25 ; Search time 47.0387 Seconds
(without alignments)
2286.133 Million cell updates/sec

Title: US-10-617-038-9
Perfect score: 1056
Sequence: 1 MIATTRDREGATMTFRURL.....IRVRNASWQHIDSLFQTOR 210

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1056	100.0	210	2 P71991	P71991 mycobacteri
2	1051	99.5	210	2 Q77ZP1	Q77ZP1 mycobacteri
3	984	93.2	198	2 Q8VJY0	Q8VJY0 mycobacteri
4	204.5	19.4	193	2 Q9R1Y8	Q9R1Y8 streptomyc
5	169.5	16.1	194	2 Q9FC26	Q9FC26 streptomyc
6	169	16.0	209	2 Q82MZ1	Q82MZ1 streptomyc
7	167	15.8	201	2 Q82N73	Q82N73 streptomyc
8	163.5	15.5	202	2 Q9K3N7	Q9K3N7 streptomyc
9	153.5	14.5	456	2 Q9ZBP2	Q9ZBP2 streptomyc
10	149.5	14.2	288	2 Q82NK2	Q82NK2 streptomyc
11	115	10.9	721	2 Q7NUF1	Q7NUF1 chromobacte
12	109	10.3	1716	2 Q96D4	Q96D4 leishmania
13	105.5	10.0	459	2 Q86622	Q86622 streptomyc
14	102.5	9.7	404	2 Q6ZJH6	Q6ZJH6 burkholderi
15	98.5	9.3	404	2 Q63T63	Q63T63 burkholderi
16	98	9.3	1577	2 Q6NVJ5	Q6NVJ5 brachydanio
17	94	8.9	1411	2 Q7JY53	Q7JY53 mycobacteri
18	92.5	8.8	256	1 US02_PRVN3	F24382 pseudorabie
19	92.5	8.8	473	2 Q9JXZ7	Q9JXZ7 neisseria m
20	91.5	8.7	635	2 Q74A83	Q74A83 geobacter s
21	90.5	8.6	426	2 Q8P9E4	Q8P9E4 xanthomonas
22	90.5	8.6	466	2 Q82PD6	Q82PD6 streptomyc
23	90.5	8.6	489	2 Q7NZE8	Q7NZE8 chromobacte
24	90.5	8.6	792	2 Q98PB0	Q98PB0 rhizobium l
25	90.5	8.6	815	2 Q9HQ54	Q9HQ54 halobacteri
26	90	8.5	297	2 Q9A677	Q9A677 caulobacter
27	90	8.5	360	2 Q7WH91	Q7WH91 bordetella
28	90	8.5	361	2 Q7VWM7	Q7VWM7 bordetella
29	90	8.5	361	2 Q7W9H1	Q7W9H1 bordetella
30	90	8.5	893	1 DAG1_MOUSE	Q62165 mus musculu
31	90	8.5	893	2 Q8BPJ7	Q8BPJ7 mus musculu

32	90	8.5	893	2 Q8CBE6	Q8CBE6 mus musculu
33	90	8.5	1400	2 Q7DA75	Q7DA75 mycobacteri
34	90	8.5	1400	2 P96419	P96419 mycobacteri
35	90	8.5	1400	2 Q7U2H8	Q7U2H8 mycobacteri
36	90	8.5	5435	2 Q9L4X2	Q9L4X2 streptomyc
37	89.5	8.5	252	2 Q9A0M1	Q9A0M1 streptococc
38	89.5	8.5	252	2 Q7CNE0	Q7CNE0 streptococc
39	89.5	8.5	337	2 Q8LNRO	Q8LNRO oryza sativ
40	89.5	8.5	393	2 Q9WX55	Q9WX55 microbacter
41	89.5	8.5	986	2 Q59506	Q59506 microbacter
42	89.5	8.5	5588	2 Q9ALM2	Q9ALM2 saccharopol
43	89	8.4	274	2 Q9HNG5	Q9HNG5 halobacteri
44	89	8.4	299	2 Q8PFG5	Q8PFG5 xanthomonas
45	89	8.4	305	2 Q8VIM2	Q8VIM2 mus musculu

ALIGNMENTS

RESULT 1

ID	P71991	PRELIMINARY;	PRT;	210 AA.
AC	P71991;			
DT	01-FEB-1997 (TRENBLrel. 02, Created)			
DT	01-FEB-1997 (TRENBLrel. 02, Last sequence update)			
DE	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN.			
GN	OrderedLocusNames=Rv1733c;			
OS	Mycobacterium tuberculosis.			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium.			
OX	NCBI_TaxID=1773;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=H37Rv;			
RX	MEDLINE=9295987; Pubmed=9634230; DOI=10.1038/31159;			
RA	Cole S.F., Brosch R., Parkhill J., Garnier T., Churcher C.M.,			
RA	Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,			
RA	Tekaia F., Badcock K., Basham D., Brown D., Chillingworth T.,			
RA	Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,			
RA	Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,			
RA	Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,			
RA	Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,			
RA	Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;			
RT	"Deciphering the biology of Mycobacterium tuberculosis from the			
RT	complete genome sequence."			
RL	Nature 393:537-544 (1998).			
DR	EMBL; BX842577; CAB03696.1; -.			
DR	PIR; H70687; H70687.			
DR	TuberculList; Rv1733c; -.			
KW	Complete proteome.			
SQ	SEQUENCE 210 AA; 22458 MW; 5D28C0B240FA9F3C CRC64;			

Query Match	100.0%;	Score 1056;	DB 2;	Length 210;
Best Local Similarity	100.0%;	Pred. No. 1.3e-79;		
Matches 210;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MIATTRDREGATMTFRURLPCRTILRVFSRNPLVRGTDRLVAVVLLAVTSLTIPFA	60	
Db	1	MIATTRDREGATMTFRURLPCRTILRVFSRNPLVRGTDRLVAVVLLAVTSLTIPFA	60	
QY	61	AAAGTAVDQSRSHVYHAQOTRHPATATVIDHEGVIDSNTTATSPPTKTIPTPARVWN	120	
Db	61	AAAGTAVDQSRSHVYHAQOTRHPATATVIDHEGVIDSNTTATSPPTKTIPTPARVWN	120	
QY	121	GIERGEVNAKPGTSGDRVGIWDSAGOLVDEPAPPARATADAALALGLWLSVAAVAG	180	
Db	121	GIERGEVNAKPGTSGDRVGIWDSAGOLVDEPAPPARATADAALALGLWLSVAAVAG	180	
QY	181	ALLALTRAILIRVRNASWQHIDSLFQTOR	210	
Db	181	ALLALTRAILIRVRNASWQHIDSLFQTOR	210	

```

RESULT 2
Q77ZP1
ID Q77ZP1 PRELIMINARY; PRT; 210 AA.
AC Q77ZP1;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE PROBABLE CONSERVED TRANSMEMBRANE PROTEIN.
GN OrderedLocusNames=Mbl1762c;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12789972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Biglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duchoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Aikin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.F., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248340; CAD94464.1; -.
GO GO:0016021; C:integral to membrane; IEA.
KW Complete proteome; Transmembrane
SQ SEQUENCE 210 AA; 22467 MW; D59F30E52524D62B6 CRC64;

Query Match 99.5%; Score 1051; DB 2; Length 210;
Best Local Similarity 99.5%; Pred. No. 3.5e-79;
Matches 209; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MIATRDREGATMTIFRLRLPCRTILRVFSRNPVGRGTDRLVAVVMLLVTVSLTTPFA 60
Db 1 MIATRDREGATMTIFRLRLPCRTILRVFSRNPVGRGTDRLVAVVMLLVTVSLTTPFA 60
Qy 61 AAAGTAVDSRSHVYAHQAQTRHPATATVIDHEGVDSNTTATSAPPRTKITVPAWVNVN 120
Db 61 AAAGTAVHDSRSHVYAHQAQTRHPATATVIDHEGVDSNTTATSAPPRTKITVPAWVNVN 120
Qy 121 GIERSEVNAKPGTKSGDRVGIVDSAGQLVDEPAPPARAIAADAAALGLWLSVAAVAG 180
Db 121 GIERSEVNAKPGTKSGDRVGIVDSAGQLVDEPAPPARAIAADAAALGLWLSVAAVAG 180
Qy 181 ALLALTRAILIRVRNASHQHDIDSLFCTOR 210
Db 181 ALLALTRAILIRVRNASHQHDIDSLFCTOR 210

RESULT 3
Q8VJY0
ID Q8VJY0 PRELIMINARY; PRT; 198 AA.
AC Q8VJY0;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MT1774;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Mayhew L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,

```

```

Fraser C.M.;
"Whole-genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains.";
J. Bacteriol. 184:5479-5490(2002).
EMBL; AE000516; AAK46047.1; -.
TIGR; MT1774; -.
Hypothetical protein.
SEQUENCE 198 AA; 21153 MW; 1CD36F20E341E91B CRC64;

Query Match 93.2%; Score 984; DB 2; Length 198;
Best Local Similarity 99.0%; Pred. No. 1.2e-73;
Matches 196; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 13 MITFRLRLPCRTILRVFSRNPVGRGTDRLVAVVMLLVTVSLTTPFAAAGTAVODSRS 72
Db 1 MITFRLRLPCRTILRVFSRNPVGRGTDRLVAVVMLLVTVSLTTPFAAAGTAVHDSRS 60
Qy 73 HVYAHQAQTRHPATATVIDHEGVDSNTTATSAPPRTKITVPAWVNVNGIERSGEVNAKP 132
Db 61 HVYAHQAQTRHPATATVIDHEGVDSNTTATSAPPRTKITVPAWVNVNGIERSGEVNAKP 120
Qy 133 GTKSGDRVGIVDSAGQLVDEPAPPARAIAADAAALGLWLSVAAVAGALLALTRAILIR 192
Db 121 GTKSGDRVGIVDSAGQLVDEPAPPARAIAADAAALGLWLSVAAVAGALLALTRAILIR 180
Qy 193 VRNASWOHDIDSLFCTOR 210
Db 181 VRNASWOHDIDSLFCTOR 198

RESULT 4
Q9RIY8
ID Q9RIY8 PRELIMINARY; PRT; 193 AA.
AC Q9RIY8;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Putative membrane protein SCJ1.26.
GN ORFNames=SCJ1.26;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
"Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939104; CAB53144.1; -.
DR PIR; T36959; T36959.
KW Complete proteome.
SQ SEQUENCE 193 AA; 20752 MW; 8C91E04DF35CE8E0 CRC64;

Query Match 19.4%; Score 204.5; DB 2; Length 193;
Best Local Similarity 32.3%; Pred. No. 4.5e-09;
Matches 63; Conservative 24; Mismatches 91; Indels 17; Gaps 5;

Qy 8 REGATMTIFRLRLPCRTILRVFSRNPVGRGTDRLVAVVMLLVTVSLTTPFAAAGTAV 67
Db 2 RAGTTT-----RRSLWRNRNPLRRREDVLEAWILLIVLVLVVALAGP---VAGVLA 49
Qy 68 QDSRSHVYAHQAQTRHPATATVIDHEGVDSNTTATSAPPRTKITVPAWVNVNGIERSG 126

```



```

Db 50 AHASGAAGAAORAEHSATATV-SDVSGSLSTGTAGRIEATV--RWTAPDGRHSG 105
QY 127 EYNAKPGTKSGDRVGVWDSAGQLVDEPAPPARAIAADAAALGLWLSVAAGALLALT 186
Db 106 KIPVDEAMKAGSRVVTWTHRNLQITAPTTTAQAGVDAAFMGAASSFAVVTAAAGYYGA 165
QY 187 RAILIRVRNASQHD 201
Db 166 RVVLNRRRRRAAWEHE 180

RESULT 5
Q9FC26 PRELIMINARY; PRT; 194 AA.
AC Q9FC26;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative integral membrane protein.
ORFNames=SC4G1.24;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieher H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieher T., Larke L., Murphy L.D., Oliver K., O'Neil S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147 (2002).
DR EMBL: AL939130; CAC01556.1; -.
KW Complete proteome.
SQ SEQUENCE 194 AA; 21759 MW; 0C794C89E4B4484D CRC64;

Query Match 16.1%; Score 169.5; DB 2; Length 194;
Best Local Similarity 30.6%; Pred. No. 3.6e-06;
Matches 56; Conservative 23; Mismatches 87; Indels 17; Gaps 7;

QY 27 RVFSRNLVRGTRLEAVVMLAVTVSLTIPFAAAAGTAV---QDSRSHVYAHQAOTR 82
Db 8 RHWRRNPLRRSDVVVEAWTKVL---VALLFVAAPALGAATAWGHGQAQAIVAEQRAER 64
QY 83 HPATATVIDH-EGVIDSNTTATSAPRTKITVPARM-VVNGIERSGEVNAKPGTKSGDRV 140
Db 65 HHVRATVDRVTGTSTGBLGQCHYRATV---HWKAPDGTGKSTTARVPADTRHGDTV 120
QY 141 GIWVDSAGQLVDEPAPPARAIAADAAALGLWLSVAAGALLA--LTRAILIRVRNASW 198
Db 121 DWLDSRGOSVPPSDSABEWQHS--ATIGSFITGTTLVLLAHRAVRANVRHMAEW 178
QY 199 QHD 201
Db 179 DRD 181

RESULT 6
Q82M21 PRELIMINARY; PRT; 209 AA.
AC Q82M21;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 26, Last annotation update)

```

```

DE Hypothetical protein.
GN OrderedLocusNames=SAV1512;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531 (2003).
DR EMBL: AP005027; BAC69222.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 209 AA; 23299 MW; 58025B6E51E72623 CRC64;

Query Match 16.0%; Score 169; DB 2; Length 209;
Best Local Similarity 32.2%; Pred. No. 4.3e-06;
Matches 56; Conservative 19; Mismatches 87; Indels 12; Gaps 6;

QY 31 RNPLVRGTRLEAVVMLAVTVSLTIPFAAAAGTAVQDSRSHVYAHQAOTRHPATATVI 90
Db 32 RNRQRSDVLEAWVLFSAWTVTLVC---GVLTGLAAHSEVQGLAREAEWRPVALVA 88
QY 91 DHEGVIDSNTTATSAPRTKITVPARM-VVNGIERSGEVNAKPGTKSGDRVGIWVDSAQ 149
Db 89 --EDAPESSAAAASGADR--VMKVRWTTANGSTHSGQARVAAGSATGTPVTWTDRCGL 144
QY 150 LVDEPAPPARAIAADAAALGLWLSVAAGALLA--LTRAILIRVRNASQHD 201
Db 145 LVTKFVPESE--AQVRGVTVGLLVGVSAAAGPFVGGRLVRGLRRRMRERWDED 196

RESULT 7
Q82N73 PRELIMINARY; PRT; 201 AA.
AC Q82N73;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative membrane protein.
GN OrderedLocusNames=SAV1430;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).
RN [2]
RP SEQUENCE FROM N.A.

```


Streptomyces avermitilis.
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Streptomyces; Streptomycetaceae; Streptomyces.
 NCBI_TaxID=33903;
 [1]
 RN
 SEQUENCE FROM N.A.
 STRAIN=NA-4680;
 MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
 Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
 Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 "Genome sequence of an industrial microorganism Streptomyces
 avermitilis: deducing the ability of producing secondary
 metabolites.";
 Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 [2]
 RN
 SEQUENCE FROM N.A.
 STRAIN=NA-4680;
 MEDLINE=22608306; PubMed=12692562;
 Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 Sakaki Y., Hattori M., Omura S.;
 "Complete genome sequence and comparative analysis of the industrial
 microorganism Streptomyces avermitilis.";
 Nat. Biotechnol. 21:526-531(2003).
 DR
 EMBL; AP005026; BAC69011.1; --
 Complete proteome; Hypothetical protein.
 SK
 SEQUENCE 288 AA; 31498 MW; 9DBBB81AAD1D10B7 CRC64;
 Query Match 14.2%; Score 149.5; DB 2; Length 288;
 Best Local Similarity 29.0%; Pred. No. 0.00025;
 Matches 54; Conservative 26; Mismatches 75; Indels 31; Gaps 9;
 QY 31 RNPVRGTRLEAVVLLAVTVSLTIPFAAAA--GTAVQDSRSHVYAHQAQTRHPATAT 88
 DB 106 RNPRLRRSDVSEAWTALVAVLLFVAAPLAGAAVGVWYDSSHSLATQRAERH-VSAT 164
 QY 89 VIDHEGVDSNTTATSAPPRTK-----ITVPAWVNGI-ERSGEVNAKPTKSGDRVG 141
 DB 165 LVEN-----APAAVPSQDSKEPLYQVKVWTEPGCGARTGPAMVPAGSLRGERTD 215
 QY 142 IWVDSAGQLVDEPAPPARAIAAAL-----AALGLWLSVAANAGALLA--LFTAILIRVN 195
 DB 216 VMLDGRGSRV--APP---VTDTLVQHVAGTSGVWTTGGVGLVLLARIARERHRM 269
 QY 196 ASWQHD 201
 DB 270 AEWEEE 275
 RESULT 11
 Q7NUF1
 ID Q7NUF1 PRELIMINARY; PRT; 721 AA.
 AC Q7NUF1;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Probable cation transport P-type ATPase (EC 3.6.3.4).
 GN OrderedLocusNames=CV2748;
 OS Chromobacterium violaceum.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Chromobacterium.
 RN NCBI_TaxID=536;
 [1]
 RN
 SEQUENCE FROM N.A.
 STRAIN=ATCC 12472 / DSM 30191;
 MEDLINE=22882880; PubMed=14500782; DOI=10.1073/pnas.1932124100;
 RA Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimaraes C.T.,
 Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
 Alves-Gomes J.A., Andrade E.M., Araripe J., de Araujo M.F.F.,
 Astolfi-Filho S., Azevedo V., Baptista A.J., Bataus L.A.M.,
 Batista J.S., Belo A., van den Berg C., Bogo M., Bonatto S.,
 Bordignon J., Brigido M.M., Brito C.A., Brocchi M., Burity H.A.,
 Camargo A.A., Cardoso D.D.P., Carneiro N.P., Carraro D.M.,

Streptomyces avermitilis.
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Streptomyceae; Streptomycetaceae; Streptomyces.
 NCBI_TaxID=33903;
 [1]
 RN
 SEQUENCE FROM N.A.
 STRAIN=NA-4680;
 MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
 Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
 Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 "Genome sequence of an industrial microorganism Streptomyces
 avermitilis: deducing the ability of producing secondary
 metabolites.";
 Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 [2]
 RN
 SEQUENCE FROM N.A.
 STRAIN=NA-4680;
 MEDLINE=22608306; PubMed=12692562;
 Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 Sakaki Y., Hattori M., Omura S.;
 "Complete genome sequence and comparative analysis of the industrial
 microorganism Streptomyces avermitilis.";
 Nat. Biotechnol. 21:526-531(2003).
 DR
 EMBL; AP005026; BAC69011.1; --
 Complete proteome; Hypothetical protein.
 SQ
 SEQUENCE 288 AA; 31498 MW; 9DBBB81AAD1D10B7 CRC64;
 Query Match 14.2%; Score 149.5; DB 2; Length 288;
 Best Local Similarity 29.0%; Pred. No. 0.00025;
 Matches 54; Conservative 26; Mismatches 75; Indels 31; Gaps 9;
 QY 31 RNPVRGTRLEAVVLLAVTVSLTIPFAAAA--GTAVQDSRSHVYAHQAQTRHPATAT 88
 DB 106 RNPRLRRSDVSEAWTALVAVLLFVAAPLAGAAVGVWYDSSHSLATQRAERHH-VSAT 164
 QY 89 VIDHEGVDSNTTATSAPPRTK-----ITVPAWVNGI-ERSGEVNAKPTKSGDRVG 141
 DB 165 LVEN-----APAAVPSVQDSKEPLYQVKVWTEPGCGARTGPAMVPAGSLRGERTD 215
 QY 142 IWVDSAGQLVDEPAPPARAIAAAL-----AALGLWLSVAANAGALLA--LFTAILIRVN 195
 DB 216 VMLDGRGSRV--APP---VTDTLVQHVAGTSGVWTTGGVGLVLLARIARERHRM 269
 QY 196 ASWQHD 201
 DB 270 AEWEEE 275
 RESULT 11
 Q7NUF1
 ID Q7NUF1 PRELIMINARY; PRT; 721 AA.
 AC Q7NUF1;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Probable cation transport P-type ATPase (EC 3.6.3.4).
 GN OrderedLocusNames=CV2748;
 OS Chromobacterium violaceum.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Chromobacterium.
 RN
 NCBI_TaxID=536;
 [1]
 RN
 SEQUENCE FROM N.A.
 STRAIN=ATCC 12472 / DSM 30191;
 MEDLINE=22882880; PubMed=14500782; DOI=10.1073/pnas.1932124100;
 RA Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimaraes C.T.,
 Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
 Alves-Gomes J.A., Andrade E.M., Araripe J., de Araujo M.F.F.,
 Astolfi-Filho S., Azevedo V., Baptista A.J., Bataus L.A.M.,
 Batista J.S., Belo A., van den Berg C., Bogo M., Bonatto S.,
 Bordignon J., Brigido M.M., Brito C.A., Brocchi M., Burity H.A.,
 Camargo A.A., Cardoso D.D.P., Carneiro N.P., Carraro D.M.,

RA Zhou L., Fraser C.M.;
RL "Structural flexibility in the Burkholderia mallei genome."
RT Proc. Natl. Acad. Sci. U.S.A. 101:14247-14251(2004).
DR EMBL; CP000010; AAU47707.1; --
KW Hypothetical protein.
SQ SEQUENCE 404 AA; 42528 MW; 6BCC25C56D0ECB9 CRC64;

Query Match 9.7%; Score 102.5; DB 2; Length 404;
Best Local Similarity 23.7%; Pred. No. 2.8;
Matches 46; Conservative 31; Mismatches 74; Indels 43; Gaps 7;

QY 46 MLLAVTVSLTIPFAAAGTAVQD-----SRSHVYAHQAQTRHPATA----- 87
DB 4 LLAAGVLSILL--SAANAAPVPSIQIQOISTAQGNWQRAQLSQVIDAYPDNARAYL 61

QY 88 --TVIDHEG-----VIDSNTTATSAPRTKITVPARWVNGIERSGEVNAKPGTKSGDRVG 141
DB 62 YGQVLDREGCPAEALAQIERAKSLDQLRFTDPSRPAQT--EARVRADARRATAAQDSRS 119

QY 142 IWDSAGQLVDEPAPPARAIAAALA-----ALGLWLSVAAVAGALLALTRAI 189
DB 120 A---TSGGMLAAPQAPQAARAFPSAAPVAPVHRGPSVGMWIGFAVLIGVIVILRKT 176

QY 190 LIRVNASWQHDID 203
DB 177 LRRARSADDDQRAAD 190

RESULT 15
Q63T63 PRELIMINARY; PRT; 404 AA.

AC Q63T63;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Putative membrane protein.
GN ORFNames=BPSL2106;
OS Burkholderia pseudomallei K96243.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=272560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K96243;
RX PubMed=15377794;
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdeno-Tarraga A.M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crosset B., Davis P., Deshazer D.,
RA Fellwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitz E., Rutherford K., Sanders M., Simmonds M.,
RA Songvilai S., Stevens K., Tumapa S., Vesaratchavest M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RT Burkholderia pseudomallei."
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR EMBL; BX571965; CAH36109.1; --
SQ SEQUENCE 404 AA; 42486 MW; 4CPA844EC9FFC433 CRC64;

Query Match 9.3%; Score 98.5; DB 2; Length 404;
Best Local Similarity 23.2%; Pred. No. 5.9;
Matches 45; Conservative 31; Mismatches 75; Indels 43; Gaps 7;

QY 46 MLLAVTVSLTIPFAAAGTAVQD-----SRSHVYAHQAQTRHPATA----- 87
DB 4 LLAAGVLSILL--SAANAAPVPSIQIQOISTAQGNWQRAQLSQVIDAYPDNARAYL 61

QY 88 --TVIDHEG-----VIDSNTTATSAPRTKITVPARWVNGIERSGEVNAKPGTKSGDRVG 141
DB 62 YGQVLDREGCPAEALAQIERAKSLDQLRFTDPSRPAQT--EARVRADARRATAAQDSRS 119

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:20:35 ; Search time 21.2674 Seconds
(without alignments)
1454.845 Million cell updates/sec

Title: US-10-617-038-10
Perfect score: 418
Sequence: 1 MTNVGQGVDAVFGVIYPPQ.....RGALFSLINELTRCAAVTG 80

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	418	100.0	80	ADI37289	Adi37289 M. tuberc
2	108	25.8	432	AAy34903	Aay34903 Chlamydia
3	104	24.9	523	ADS30337	Ads30337 Bacterial
4	101	24.2	432	AAy36947	Aay36947 Amino aci
5	100	23.9	452	ADS25703	Ads25703 Bacterial
6	100	23.9	452	ADS26186	Ads26186 Bacterial
7	100	23.9	452	ADS22660	Ads22660 Bacterial
8	97	23.2	428	ADS27881	Ads27881 Bacterial
9	96	23.0	525	ADA36480	Ada36480 Acinetoba
10	95	22.7	427	ADS1688	Ads1688 Bacterial
11	95	22.7	445	ADN25654	Adn25654 Bacterial
12	94	22.5	463	ADS42235	Ads42235 Bacterial
13	88.5	21.2	417	ADS43304	Ads43304 Bacterial
14	88	21.1	412	ADS44485	Ads44485 Bacterial
15	88	21.1	454	ADS22965	Ads22965 Bacterial
16	87	20.8	491	ADS28762	Ads28762 Bacterial
17	86.5	20.7	441	ADN25351	Adn25351 Bacterial
18	86	20.6	420	ADL05645	Adl05645 M. catarr
19	86	20.6	453	ADS22789	Ads22789 Bacterial
20	85	20.3	402	AAU38461	Aau38461 Salmonell
21	83	19.9	436	ADS28114	Ads28114 Bacterial
22	83	19.9	455	ABU44607	Abu44607 Protein e
23	82.5	19.7	577	3 AAG41842	Aag41842 Arabidops
24	82.5	19.7	637	3 AAG41841	Aag41841 Arabidops
25	82.5	19.7	683	3 AAG41840	Aag41840 Arabidops

ALIGNMENTS

RESULT 1
ADI37289
ID ADI37289 standard; protein; 80 AA.

XX AC ADI37289;
XX DT 22-APR-2004 (first entry)
XX DE M. tuberculosis low oxygen induced antigen Rv1734c SEQ ID NO:10.
XX KW mycobacterial infection; vaccine; tuberculosis;
KW Mycobacterium tuberculosis; immunisation; antibacterial; gene therapy;
KW low oxygen induced antigen.

XX OS Mycobacterium tuberculosis.
XX PN WO2004006952-A2.
XX PD 22-JAN-2004.

XX PF 08-JUL-2003; 2003WO-DK000477.

XX PR 13-JUL-2002; 2002DK-00001098.
(STAT-) STATENS SERUM INST.

XX PI Andersen P, Rosenkrands I, Stryhn A;
XX DR WPI; 2004-122778/12.

XX PT N-PSDB; ADI37334.
Use of one or more polypeptides or their fragments, which are expressed during the latent stage of the mycobacterial infection, and/or nucleic acids encoding the polypeptides, for a therapeutic vaccine against tuberculosis.

XX PS Claim 3; SEQ ID NO 10; 76pp; English.

XX CC The present invention describes polypeptides or their fragments, which are expressed during the latent stage of a mycobacterial infection, and/or nucleic acids encoding the polypeptides, which are useful for creating a therapeutic vaccine against tuberculosis. Also described: (1) a therapeutic vaccine against tuberculosis comprising one or more polypeptides; (2) a method for treating an animal, including a human being, with tuberculosis caused by virulent mycobacteria, e.g. by Mycobacterium tuberculosis, M. africanum or M. bovis; (3) a method for immunising an animal, including a human being, against tuberculosis caused by virulent mycobacteria; (4) a method of diagnosing tuberculosis

AbO77262 Pseudomon
Abu36285 Protein e
Aau34503 E. coli c
Abb99316 Amino aci
Aao19317 Escherich
Abg70593 E. coli L
Abg71288 Escherich
Ada08241 Escherich
Adn18106 Bacterial
Abo66743 Klebsiell
Adn26271 Bacterial
Adf05092 Bacterial
Aag12342 Zea mays
Aag12341 Zea mays
Ads28688 Bacterial
Adn17745 Bacterial
Ads42753 Bacterial
Abm67888 Photorhab
Ads4838 Bacterial
Abu17519 Protein e

CC caused by virulent mycobacteria in an animal, including a human being;
CC (5) a method for diagnosing previous or ongoing infection with a virulent
CC mycobacterium; and (6) a method of diagnosing Mycobacterium tuberculosis
CC infection in a subject. The polypeptides have antibacterial activities,
CC and can be used in vaccines and in gene therapy. The polypeptides are
CC useful for the manufacture of a therapeutic vaccine for treating an
CC individual who is infected by a virulent mycobacterium, e.g. M.
CC tuberculosis, and who is not vaccinated with BCG against tuberculosis.
CC The present sequence represents a low oxygen induced antigen, which is
CC used in the exemplification of the present invention.

XX
SQ Sequence 80 AA;
Query Match 100.0%; Score 418; DB 8; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.3e-48; Mismatches 0; Indels 0; Gaps 0;
Matches 80; Conservative 0;
QY 1 MTNVGDQGVDAVFGVIYPPQVALVSFGKPAQRVCAVDGAIHVMTTVLATLPADHGCSDH 60
Db 1 MTNVGDQGVDAVFGVIYPPQVALVSFGKPAQRVCAVDGAIHVMTTVLATLPADHGCSDH 60
QY 61 RGALFFLSINELTRCAAVTG 80
Db 61 RGALFFLSINELTRCAAVTG 80

RESULT 2
AAV34903
ID AAV34903 standard; protein; 432 AA.
XX
AC AAV34903;
DT 17-OCT-2003 (revised)
DT 13-SEP-1999 (first entry)
XX
DE Chlamydia pneumoniae transmembrane protein sequence.
XX
KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis; vaccine;
KW neutralising epitope.

XX
OS Chlamydia pneumoniae.
XX
PN WO927105-A2.
XX
PD 03-JUN-1999.
PF 20-NOV-1998; 98WO-IB001890.
XX
XX 21-NOV-1997; 97PR-00014673.
PR 04-NOV-1998; 98US-0107078P.
XX
PA (GEST) GENSET.

XX
PI Griffiths R;
XX
XX WPI; 1999-357842/30.
XX
PT Genome sequence of Chlamydia pneumoniae.
XX
PS Page 838; Disclosure; 1912pp; English.

XX
CC AAV34584-Y35879 represent the proteins encoded by all the open reading
CC frames in the complete genome (see AAX91990) of Chlamydia pneumoniae. C.
CC pneumoniae causes respiratory disease such as pneumonia and bronchitis
CC and is thought to be a contributing factor in heart disease, sarcoidosis,
CC sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The
CC polypeptides encoded by the open reading frames of the C. pneumoniae
CC genome (see AAV34584-Y35879) can be used in immunogenic compositions as
CC vaccines. Vectors containing C. pneumoniae nucleotide sequences can also
CC be used as immunogenic compositions, especially where the vector directs
CC the expression of a neutralising epitope of C. pneumoniae. (Updated on 17
CC -OCT-2003 to standardise OS field)

XX
SQ Sequence 432 AA;
Query Match 25.8%; Score 108; DB 2; Length 432;
Best Local Similarity 26.9%; Pred. No. 1.6e-05;
Matches 21; Conservative 19; Mismatches 38; Indels 0; Gaps 0;
QY 1 MTNVGDQGVDAVFGVIYPPQVALVSFGKPAQRVCAVDGAIHVMTTVLATLPADHGCSDH 60
Db 351 VSNLGMTITFTTAIVNPPQAILAVGSVTQOALVDGEITGTSCNLTLSVDHIVDGY 410
QY 61 RGALFFLSINELTRCAAV 78
Db 411 PAAMFMKLEKLEAPAV 428

RESULT 3
ADS30337
ID ADS30337 standard; protein; 523 AA.
XX
AC ADS30337;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #19370.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
XX WPI; 2004-061375/06.
XX

PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.

XX
PS Claim 1; SEQ ID NO 19370; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,

CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 523 AA;

Query Match 24.9%; Score 104; DB 8; Length 523;
 Best Local Similarity 31.9%; Pred. No. 6.9e-05;
 Matches 23; Conservative 14; Mismatches 35; Indels 0; Gaps 0;
 QY 1 MTNVGQGVDAVFGVTPPOVALVSGKPAQRVCAVDGAHVMTTTLATLPADHGSDDH 60
 Db 443 VSNLGMYGIDQFVAVINPPQASILAVGAVSEKAVVRDQLAVRKMNTATLSCDHRVIDGA 502
 QY 61 RGALFFLSINEL 72
 Db 503 IGAEFLRELGL 514

RESULT 4
 AAY36947
 ID AAY36947 standard; protein; 432 AA.
 AC AAY36947;
 XX
 DT 07-OCT-1999 (first entry)
 XX
 DE Amino acid sequence of a Chlamydia trachomatis protein.
 XX
 KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
 KW paratrachoma; inclusion conjunctivitis; genital disease; perihhepatitis;
 KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
 KW Bartholinitis; pneumopathy; venereal lymphogranulomatosis.
 XX
 OS Chlamydia trachomatis.
 XX
 PN WO9928475-A2.
 XX
 PD 10-JUN-1999.
 XX
 PF 27-NOV-1998; 98WO-18001939.
 XX
 PR 28-NOV-1997; 97FR-00015041.
 PR 17-DEC-1997; 97FR-00016034.
 PR 04-NOV-1998; 98US-0107077P.
 XX
 PA (GEST) GENSET.
 XX
 PI Griffiths R;
 XX
 DR WPI; 1999-371125/31.
 XX
 PT Genome sequence of Chlamydia trachomatis.
 XX
 PS Disclosure; Page 797; 1755pp; English.
 XX
 CC AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
 CC of Chlamydia trachomatis (see AAZ01425). The polypeptides can be used as
 CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
 CC can also be used to control growth of the microorganism. Chlamydia
 CC trachomatis is responsible for a large number of diseases, e.g. eye
 CC diseases such as conventional trachoma, nonendemic trachoma,
 CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
 CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
 CC perihhepatitis, Bartholinitis; pneumopathy in breast feeding infants; and

CC venereal lymphogranulomatosis. The polypeptides of the invention may be
 CC of use in treating these diseases
 XX
 SQ Sequence 432 AA;
 Query Match 24.2%; Score 101; DB 2; Length 432;
 Best Local Similarity 25.6%; Pred. No. 0.00014;
 Matches 20; Conservative 20; Mismatches 38; Indels 0; Gaps 0;
 QY 1 MTNVGQGVDAVFGVTPPOVALVSGKPAQRVCAVDGAHVMTTTLATLPADHGSDDH 60
 Db 351 VSNLGMTGISDFTAILNPPQAILAVGSVEEPVVLNGELAVGSCMLTSLVDHRVIDGY 410
 QY 61 RGALFFLSINELTRCAAV 78
 Db 411 PAAMFMKRLQRLLEAPSV 428
 RESULT 5
 ADS25703
 ID ADS25703 standard; protein; 452 AA.
 XX
 AC ADS25703;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polypeptide #14736.
 XX
 KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX
 OS Bacteria.
 XX
 PN US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 DR WPI; 2004-061375/06.
 XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 14736; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with

CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 452 AA;

Query Match 23.9%; Score 100; DB 8; Length 452;
Best Local Similarity 33.3%; Pred. No. 0.0002;
Matches 21; Conservative 10; Mismatches 32; Indels 0; Gaps 0;

QY 1 MTNVGDQGVDAVFGVIYPPQVALVSFGKPAQRVCAVDGAIHVMTVTVLATLPADHCCSDDH 60
Db 372 VSNMGMGVKSFSAVINPPHATILAVGAGEQRAVVKNGEIKIANVMTVTLTSTDHRCVDGA 431
QY 61 RGA 63
Db 432 LGA 434

RESULT 6
ADS26186
ID ADS26186 standard; protein; 452 AA.
XX ADS26186;
XX
DT 02-DEC-2004 (first entry)
DE Bacterial polypeptide #15219.
XX

KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

XX Bacteria.
XX US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX
DR WPI; 2004-061375/06.
XX

PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.

PS Claim 1; SEQ ID NO 15219; 122pp; English.

CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 452 AA;

Query Match 23.9%; Score 100; DB 8; Length 452;
Best Local Similarity 33.3%; Pred. No. 0.0002;
Matches 21; Conservative 10; Mismatches 32; Indels 0; Gaps 0;

QY 1 MTNVGDQGVDAVFGVIYPPQVALVSFGKPAQRVCAVDGAIHVMTVTVLATLPADHCCSDDH 60
Db 372 VSNMGMGVKSFSAVINPPHATILAVGAGEQRAVVKNGEIKIANVMTVTLTSTDHRCVDGA 431
QY 61 RGA 63
Db 432 LGA 434

RESULT 7
ADS22660
ID ADS22660 standard; protein; 452 AA.

XX ADS22660;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #11693.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

XX Bacteria.
XX US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX

KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX Bacteria.
OS US2003233675-A1.
XX
PN 18-DEC-2003.
XX
PD 20-FEB-2003; 2003US-00369493.
XX
PF 21-FEB-2002; 2002US-0360039P.
XX
PR (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
DR New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 8307; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 445 AA;

Query Match 22.7%; Score 95; DB 8; Length 445;
Best Local Similarity 31.9%; Pred. No. 0.00095;
Matches 23; Conservative 12; Mismatches 37; Indels 0; Gaps 0;

Qy 1 MTNVGOGVDAVEGVYPPQVALVSFGKPAQRCAVDGAHVTWTVLATIPADHGCSDDH 60
Db 365 VSNLGMFGVSFSAVINPPEAAILAVGAMQEPFVRDGRIVARHTALELSVDHRAVDGA 424

Qy 61 RGALFFFLSINEL 72
Db 425 VGRAFLKDLAEV 436

RESULT 12
ADS42235
ID ADS42235 standard; protein; 463 AA.
XX
AC ADS42235;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #20665.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX Bacteria.
OS US2003233675-A1.
XX
PN 18-DEC-2003.
PD 20-FEB-2003; 2003US-00369493.
XX
PF 21-FEB-2002; 2002US-0360039P.
XX
PR (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
DR New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 20665; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 463 AA;

Query Match 22.5%; Score 94; DB 8; Length 463;
Best Local Similarity 34.9%; Pred. No. 0.0014;

CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 412 AA;
Query Match 21.1%; Score 88; DB 8; Length 412;
Best Local Similarity 30.6%; Pred. No. 0.0077;
Matches 22; Conservative 10; Mismatches 40; Indels 0; Gaps 0;
QY 1 MTNVGQGVDAVFGVIYPPQVALVSFGKPAQRCVCAVDGAIHVTTLATLPADHGCSDDH 60
Db 326 ITNVGSIGSVIGPVIYPPNVAIIGVHRLVERPVYVDGELKPKRIGFVSLSPDHRALEGA 385
QY 61 RGALFFLSINEL 72
Db 386 YATRFLEMEVKRL 397
RESULT 15
ADS22965
ID ADS22965 standard; protein; 454 AA.
XX
AC ADS22965;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #11998.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 19-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a

PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 11998; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 454 AA;
Query Match 21.1%; Score 88; DB 8; Length 454;
Best Local Similarity 27.8%; Pred. No. 0.0087;
Matches 20; Conservative 13; Mismatches 39; Indels 0; Gaps 0;
QY 1 MTNVGQGVDAVFGVIYPPQVALVSFGKPAQRCVCAVDGAIHVTTLATLPADHGCSDDH 60
Db 374 VSNLGMFGIKDFAAVINPPHATILAVGAGEERAVVKNGEIKIATVNSVTLDHRAVDGA 433
QY 61 RGALFFLSINEL 72
Db 434 LGAEELLVAFKRL 445
Search completed: October 15, 2005, 04:26:27
Job time : 22.2674 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 01:42:44 ; Search time 6.56835 Seconds
(without alignments)
909.197 Million cell updates/sec

Title: US-10-617-038-10
Perfect score: 418
Sequence: 1 MTNVGQGVDAVFGVIYPPQ.....RGALFFLSINELTRCAAVTG 80

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCUTUS_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	25.8	432	4	US-09-198-452A-321
2	108	25.8	432	4	US-09-438-185A-308
3	104	24.9	514	4	US-09-902-540-10091
4	96	23.0	525	4	US-09-328-352-7767
5	86	20.6	420	4	US-09-540-236-3331
6	82	19.6	423	4	US-09-252-991A-26008
7	80	19.1	405	1	US-08-370-193A-9
8	80	19.1	405	4	US-10-078-107-6
9	80	19.1	405	4	US-10-077-751-6
10	80	19.1	407	3	US-09-271-438A-11
11	80	19.1	419	4	US-09-489-039A-13260
12	79	18.9	403	4	US-09-543-681A-5377
13	74.5	17.8	168	3	US-08-975-762-37
14	74.5	17.8	168	3	US-08-821-324-37
15	74.5	17.8	168	3	US-09-295-028-37
16	74.5	17.8	168	3	US-09-106-582-37
17	74.5	17.8	168	4	US-09-159-469-37
18	74.5	17.8	168	4	US-09-693-542-37
19	74.5	17.8	340	3	US-08-975-762-54
20	74.5	17.8	340	3	US-09-295-028-54
21	74.5	17.8	340	3	US-09-106-582-54
22	74.5	17.8	340	4	US-09-159-469-54
23	74.5	17.8	340	4	US-09-693-542-54
24	74.5	17.8	406	3	US-09-066-046-21
25	74	17.7	345	4	US-09-540-226-3797
26	74	17.7	405	4	US-09-328-352-4239
27	73	17.5	958	1	US-08-426-236-4

28 72.5 17.3 433 4 US-09-710-279-1332 Sequence 1332, Ap
29 72.5 17.3 437 3 US-09-134-001C-3418 Sequence 3418, Ap
30 72.5 17.3 480 3 US-09-108-020-6 Sequence 6, Appli
31 72.5 17.3 480 4 US-09-685-296-6 Sequence 6, Appli
32 72 17.2 422 1 US-08-403-545-4 Sequence 4, Appli
33 72 17.2 422 3 US-08-404-381-4 Sequence 4, Appli
34 71 17.0 375 4 US-09-583-110-3837 Sequence 3837, Ap
35 71 17.0 377 4 US-09-107-433-3088 Sequence 3088, Ap
36 71 17.0 407 3 US-09-271-438A-4 Sequence 10, Appli
37 71 17.0 407 3 US-09-271-438A-10 Sequence 2, Appli
38 71 17.0 407 4 US-10-078-107-2 Sequence 2, Appli
39 71 17.0 407 4 US-10-077-751-2 Sequence 393, App
40 69 16.5 147 4 US-09-198-452A-393 Sequence 379, App
41 69 16.5 365 4 US-09-438-185A-379 Sequence 565, App
42 68.5 16.4 393 4 US-09-198-452A-565 Sequence 529, App
43 68.5 16.4 395 4 US-09-438-185A-529 Sequence 58510, A
44 68 16.3 138 4 US-09-270-767-58510 Sequence 43172, A
45 68 16.3 433 4 US-09-270-767-43172

ALIGNMENTS

RESULT 1
US-09-198-452A-321
; Sequence 321, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 321
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-321

Query Match 25.8%; Score 108; DB 4; Length 432;
Best Local Similarity 26.9%; Pred. No. 1.9e-06;
Matches 21; Conservative 19; Mismatches 38; Indels 0; Gaps 0;

QY 1 MTNVGQGVDAVFGVIYPPQALVSGKPAORCAVDAIHMVMTTLATLPADHGCDDH 60
DB 351 VSNLGTGTEFTAIYNPFOAILVSGVTEQALVLDGEITGTCNLTLSVDHRVIDGY 410
QY 61 RGALFFLSINELTRCAAV 78
DB 411 PAAMFMKRLQKILEAPAV 428

RESULT 2
US-09-438-185A-308
; Sequence 308, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08

Qy	12	VFG-----VIYPPQVALVSFKPAQRVCANDGAIHWMTTVLATLPADHCSDDDHGRGALF	55
Db	346	FGSLSTPIVNPPTAILGMHKIQERPMVNGQVILEMMLYALSVDHLIDGKEAVSF	405
Qy	66	FLSINEL	72
		:::	
Db	406	LVAIKDL	412

RESULT 7
US-08-370-193A-9
; Sequence 9, Application US/08370193A
; Patent No. 5573945
; GENERAL INFORMATION:
; APPLICANT: ONO, EIJI
; APPLICANT: TSUJIMOTO, NOBUHARU
; APPLICANT: MATSUI, KAZUHIKO
; APPLICANT: KURAHASHI, KAZUHIKO
; TITLE OF INVENTION: MUTANT AND METHOD FOR PRODUCING
; L-GLUTAMIC ACID BY FERMENTATION
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCELLELAND, MAIER & NEUSTADT,
; ADDRESSEE: P. C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202

```
Query Match      19.1%; Score 80; DB 1; Length 405;
Best Local Similarity 29.5%; Pred.No. 0.014;
Matches 23; Conservative 11; Mismatches 32; Indels 12; Gaps 2;
```

Qy 1 MTNVGDOGVDVAFG-----VIYPPQVALVSFGPAORVCADVGAIHVMVTVLATLPADH 54
||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
323 ITNGG-----VFGLMSTPIINPFPQAILGMHAIKDRPMVAVGQVEILPMYLALSVDH 376

Qy 55 GCSDDHKGALFFLSINEL 72
||| |::|||

Db 377 RLIDGRESVGFLVTIKEL 394

RESULT 8
US-10-078-107-6
; Sequence 6, Application US/10078107
; Patent No. 6596517
; GENERAL INFORMATION:
; APPLICANT: IZUI, HIROSHI

```

; APPLICANT: HARA, YOSHIHIKO
; APPLICANT: SATO, MASAKAZU
; APPLICANT: AKIYOSHI, NAOKI
; TITLE OF INVENTION: METHOD FOR PRODUCING L-GLUTAMIC ACID
; FILE REFERENCE: 2198460SUO
; CURRENT APPLICATION NUMBER: US/10/078,107
; CURRENT FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: JP 2001-044134
; PRIOR FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-078-107-6

```

```
Query Match      19.1%; Score 80; DB 4; Length 405;
Best Local Similarity 29.5%; Pred.No. 0.014;
Matches 23; Conservative 11; Mismatches 32; Indels 12; Gaps 2;

Qy   1 MTNVGQGVDVAFG-----VIYPQVALVSFGKPAQRCAVGDGAIHVMTTVLATLPADH 54
      ||| ||| :||| :| | :| | :| | :| | :| |
Db   323 ITNGG-----VGSLMSTPIINPQSALIGMHAIKDRPMAVNGQEILPMMYLALSVDH 376

Qy   55 GCSDDDRHALFFLSINEL 72
      | :| | |
Db   377 RLIDGRESVGFLVTIKEL 394
```

```

RESULT 9
US-10-077-751-6
; Sequence 6, Application US/10077751
; Patent No. 6653110
; GENERAL INFORMATION:
; APPLICANT: SATO, MASAKAZU
; APPLICANT: AKIYOSHI, NAKOI
; TITLE OF INVENTION: METHOD FOR PRODUCING L-GLUTAMIC ACID
; FILE REFERENCE: 219849US0
; CURRENT APPLICATION NUMBER: US/10/077,751
; CURRENT FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: JP 2001-044135
; PRIOR FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 405
; TYPE: prt
; ORGANISM: Escherichia coli
US-10-077-751-6

```

Query Match	19.1%	Score 80;	DB 4;	Length 405;
Best Local Similarity	29.5%;	Pred.No.	0.014;	
Matches	23;	Conservative	11;	Mismatches 32; Indels 12; Gaps 2;
Qy	1	MTNVGQGVDVAVFG-----VIYPQVALVSFKPAQRVCACVGAIHVMTTVLATLPADH	54	
	:	:::	:::	:::
Db	323	ITNGG-----VFGLSMSTPIINPPOSAILGMHAIKDRPMVAVGVEILPMMYLALSYDH	376	
	:	:::	:::	:::
Qy	55	GCSDDHRGALFFLSINEL	72	
		:::	:::	
Db	377	RLIDGRESVGFLVTIKEL	394	
		:::	:::	

```

RESULT 10
; US-09-271-438A-11
; Sequence 11, Application US/09271438A
; Patent No. 6331419
; GENERAL INFORMATION:
; APPLICANT: IZUI, Hiroshi
; APPLICANT: ONO, Eiichi
; APPLICANT: MATSUI, Kazuhiko
; APPLICANT: MORIYA, Mika
;

```

```
; APPLICANT: ITO, Hisao
; APPLICANT: HARA, Yoshihiko
; TITLE OF INVENTION: L-GLUTAMIC ACID-PRODUCING BACTERIUM AND METHOD FOR PRODUCING L-GLUTAMIC ACID
; FILE REFERENCE: 0010-0989-0
; CURRENT APPLICATION NUMBER: US/09/271,438A
; CURRENT FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: JP10-69068
; PRIOR FILING DATE: 1998-03-18
; PRIOR APPLICATION NUMBER: JP10-297129
; PRIOR FILING DATE: 1998-10-19
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-271-438A-11

Query Match          19.1%; Score 80; DB 3; Length 407;
Best Local Similarity 29.5%; Pred. No. 0.014; 32; Indels 12; Gaps 2;
Matches 23; Conservative 11; Mismatches 11;

QY 1 MTNVGDSQGVDAVFG-----VIYPPQVALVSFGKPAQRVCAVDGAIHVMTTVLTATLPADH 54
Db 325 ITNGG-----VFGSLMSTPIINPPQSAILGHWAIKDRPMVAVNGQVEILPMYLAISYDH 378

QY 55 GCSDDHRGALFSLNEL 72
Db 379 RLIDGRESVGLVTKEL 396

RESULT 11
US-09-489-039A-13260
; Sequence 13260, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13260
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13260

Query Match          19.1%; Score 80; DB 4; Length 419;
Best Local Similarity 29.5%; Pred. No. 0.014;
Matches 23; Conservative 11; Mismatches 32; Indels 12; Gaps 2;

QY 1 MTNVGDSQGVDAVFG-----VIYPPQVALVSFGKPAQRVCAVDGAIHVMTTVLTATLPADH 54
Db 337 ITNGG-----VFGSLMSTPIINPPQSAILGHWAIKDRPMVAVNGQVEILPMYLAISYDH 390

QY 55 GCSDDHRGALFSLNEL 72
Db 391 RLIDGRESVGLVAIKEL 408

RESULT 12
US-09-543-681A-5377
; Sequence 5377, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS FOR DIAGNOSTICS AND THERAPEUTICS
```

```
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5377
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5377

Query Match          18.9%; Score 79; DB 4; Length 403;
Best Local Similarity 28.2%; Pred. No. 0.019;
Matches 22; Conservative 12; Mismatches 32; Indels 12; Gaps 2;

QY 1 MTNVGDSQGVDAVFG-----VIYPPQVALVSFGKPAQRVCAVDGAIHVMTTVLTATLPADH 54
Db 321 ITNGG-----VFGSLMSTPIINPPQSAILGHWAIKDRPMVAVNGQVEILPMYLAISYDH 374

QY 55 GCSDDHRGALFSLNEL 72
Db 375 RLIDGRESVGLVTKEM 392

RESULT 13
US-08-975-762-37
; Sequence 37, Application US/08975762
; Patent No. 6207169
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF BACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,762
; FILING DATE: 21-MAR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.439
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; TELEFAX: 206-682-6031
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 168 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Ehrlichia
US-08-975-762-37

Query Match          17.8%; Score 74.5; DB 3; Length 168;
Best Local Similarity 32.8%; Pred. No. 0.025; 35; Indels 1; Gaps 1;
Matches 22; Conservative 9; Mismatches 35; Indels 1; Gaps 1;
```

QY 1 MTNVGQGVDAVFGVIYPPQVALVSFGKPAQRCAVDGAIHVMTTVLATLPADHGCSDDH 60
Db 86 ITNGGVYGSLLSTPIINPPQSGILGMHAIQORPVAVDGKVEIRPMMYLALSYDHRIVDG- 144
QY 61 RGALFFEL 67
Db 145 QGAVTFL 151

RESULT 14

US-08-821-324-37
; Sequence 37, Application US/08821324
; Patent No. 6231869
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND
; THERAPY
; NUMBER OF SEQUENCES: 38

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/821,324
; FILING DATE: 21-MAR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:

; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.439
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; TELEFAX: 206-682-6031

; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 168 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Ehrlichia

US-08-821-324-37

Query Match 17.8%; Score 74.5; DB 3; Length 168;
Best Local Similarity 32.8%; Pred. No. 0.025;
Matches 22; Conservative 9; Mismatches 35; Indels 1; Gaps 1;

QY 1 MTNVGQGVDAVFGVIYPPQVALVSFGKPAQRCAVDGAIHVMTTVLATLPADHGCSDDH 60
Db 86 ITNGGVYGSLLSTPIINPPQSGILGMHAIQORPVAVDGKVEIRPMMYLALSYDHRIVDG- 144
QY 61 RGALFFEL 67
Db 145 QGAVTFL 151

RESULT 15

US-09-295-028-37
; Sequence 37, Application US/09295028
; Patent No. 6277381
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.

; APPLICANT: Houghton, Raymond L.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
; TITLE OF INVENTION: AND TREATMENT OF EHRlichia INFECTION
; FILE REFERENCE: 210121.439C4
; CURRENT APPLICATION NUMBER: US/09/295,028
; CURRENT FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 37
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Ehrlichia sp.
US-09-295-028-37

Query Match 17.8%; Score 74.5; DB 3; Length 168;
Best Local Similarity 32.8%; Pred. No. 0.025;
Matches 22; Conservative 9; Mismatches 35; Indels 1; Gaps 1;

QY 1 MTNVGQGVDAVFGVIYPPQVALVSFGKPAQRCAVDGAIHVMTTVLATLPADHGCSDDH 60
Db 86 ITNGGVYGSLLSTPIINPPQSGILGMHAIQORPVAVDGKVEIRPMMYLALSYDHRIVDG- 144
QY 61 RGALFFEL 67
Db 145 QGAVTFL 151

Search completed: October 15, 2005, 03:24:16
Job time : 7.56835 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:14:10 ; Search time 19.1311 Seconds
(without alignments)
1742.576 Million cell updates/sec

Title: US-10-617-038-10

Perfect score: 418
Sequence: 1 MTNVGQGVDAVFGVYIPFQ.....RGALFFLSINELTRCAAVTG 80

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	418	100.0	80	15	US-10-617-038-10
2	108	25.8	432	15	Sequence 10, Appl
3	104	24.9	523	15	Sequence 321, App
4	100	23.9	452	15	Sequence 19370, A
5	100	23.9	452	15	Sequence 11693, A
6	100	23.9	452	15	Sequence 14736, A
7	97	23.2	428	15	Sequence 15219, A
8	95	22.7	427	15	Sequence 16914, A
9	95	22.7	445	15	Sequence 10721, A
10	94	22.5	463	15	Sequence 8307, Ap
11	89	21.3	489	17	Sequence 20665, A
					Sequence 1, Appl

12	88.5	21.2	417	15	US-10-369-493-21734	Sequence 21734, A
13	88	21.1	412	15	US-10-369-493-22915	Sequence 22915, A
14	88	21.1	454	15	US-10-369-493-11998	Sequence 11998, A
15	87	20.8	491	15	US-10-369-493-17795	Sequence 17795, A
16	86.5	20.7	441	15	US-10-369-493-8004	Sequence 8004, Ap
17	86	20.6	453	15	US-10-369-493-11822	Sequence 11822, A
18	85	20.3	402	9	US-09-815-242-14054	Sequence 14054, A
19	83	19.9	436	15	US-10-369-493-17147	Sequence 17147, A
20	83	19.9	455	15	US-10-282-122A-72531	Sequence 72531, A
21	81.5	19.5	484	16	US-10-437-963-144338	Sequence 144338, A
22	81	19.4	402	15	US-10-282-122A-64209	Sequence 64209, A
23	80	19.1	405	9	US-09-815-242-10096	Sequence 10096, A
24	80	19.1	405	13	US-10-078-107-6	Sequence 6, Appli
25	80	19.1	405	13	US-10-077-751-6	Sequence 6, Appli
26	80	19.1	405	14	US-10-077-745-6	Sequence 6, Appli
27	80	19.1	405	14	US-10-338-915-6	Sequence 6, Appli
28	80	19.1	405	15	US-10-369-493-759	Sequence 759, App
29	80	19.1	407	14	US-10-315-023-11	Sequence 11, Appl
30	79.5	19.0	461	15	US-10-369-493-8924	Sequence 8924, Ap
31	78.5	18.8	442	15	US-10-369-493-17721	Sequence 17721, A
32	78.5	18.8	543	16	US-10-425-115-339881	Sequence 339881, A
33	78.5	18.8	543	16	US-10-425-115-339903	Sequence 339903, A
34	78.5	18.8	548	16	US-10-437-963-162471	Sequence 162471, A
35	78	18.7	403	15	US-10-369-493-398	Sequence 398, App
36	78	18.7	404	15	US-10-369-493-21183	Sequence 21183, A
37	77.5	18.5	108	16	US-10-767-701-42485	Sequence 42485, A
38	77.5	18.5	200	16	US-10-425-115-339892	Sequence 339892, A
39	77.5	18.5	242	15	US-10-425-114-52337	Sequence 52337, A
40	77.5	18.5	342	15	US-10-425-114-41975	Sequence 41975, A
41	77.5	18.5	458	16	US-10-425-115-223762	Sequence 223762, A
42	77.5	18.5	488	15	US-10-425-114-66723	Sequence 66723, A
43	77	18.4	407	15	US-10-369-493-13871	Sequence 13871, A
44	75.5	18.1	419	15	US-10-282-122A-45443	Sequence 45443, A
45	75.5	18.1	429	15	US-10-369-493-16532	Sequence 16532, A

ALIGNMENTS

RESULT 1

US-10-617-038-10
; Sequence 10, Application US/10617038
; Publication No. US20040057963A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Peter
; APPLICANT: Rosenkrands, Ida
; APPLICANT: Stryhn, Anette
; TITLE OF INVENTION: Therapeutic TB Vaccine
; FILE REFERENCE: SS15AUSA
; CURRENT APPLICATION NUMBER: US/10/617,038
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: DK PA 2002 01098
; PRIOR FILING DATE: 2002-07-13
; PRIOR APPLICATION NUMBER: US 60/401,725
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-617-038-10

Query Match	100.0%	Score	418	DB	15	Length	80
Best Local Similarity	100.0%	Pred. No.	1.8e-46				
Matches	80	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MTNVGQGVDAVFGVYIPFQVALVSGKPAQRCAVDAIHVMTTVLATLPADHGSDDH	60				
Db	1	MTNVGQGVDAVFGVYIPFQVALVSGKPAQRCAVDAIHVMTTVLATLPADHGSDDH	60				
Qy	61	RGALFFLSINELTRCAAVTG	80				


```

RESULT 8
US-10-369-493-10721
; Sequence 10721, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 10721
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Sphingomonas aromaticivorans
US-10-369-493-10721

Query Match          22.7%; Score 95; DB 15; Length 427;
Best Local Similarity 34.9%; Pred. No. 0.002;
Matches 22; Conservative 11; Mismatches 30; Indels 0; Gaps 0;

QY 1 MTNVGQQGVDAVFGVIYPPQVALVSGFKPAQRVCAVDGAIHVMTTVLATLPADHGCSDDH 60
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 347 LSNLGMFGIKNFDVINPPQAMAVGAEQRPYVIDGALGIATVMSATSGSFDHRAIDGA 406
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

QY 61 RGA 63
   ||
Db 407 DGA 409

RESULT 9
US-10-369-493-8307
; Sequence 8307, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 8307
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Thermobifida fusca
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(445)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-8307

Query Match          22.7%; Score 95; DB 15; Length 445;
Best Local Similarity 31.9%; Pred. No. 0.0021;
Matches 23; Conservative 12; Mismatches 37; Indels 0; Gaps 0;

QY 1 MTNVGQQGVDAVFGVIYPPQVALVSGFKPAQRVCAVDGAIHVMTTVLATLPADHGCSDDH 60
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 365 VSNLGMFGVGSFVNPPEAAILAVGAMOEPPVVRDCEIVARHTIALELSVDHRAVDGA 424
   ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

```

```
QY 61 RGALFFLSINEL 72
Db 425 VQAFLKDLAEV 436

RESULT 10
US-10-369-493-20665
; Sequence 20665, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 20665
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Rhodospseudomonas palustris
US-10-369-493-20665

Query Match 22.5%; Score 94; DB 15; Length 463;
Best Local Similarity 34.9%; Pred. No. 0.0029;
Matches 22; Conservative 8; Mismatches 33; Indels 0; Gaps 0;

QY 1 MTNVGQGVDAVFGVIYPPQVALVSFGKPAQRVCVAVDGAHVMTTVLATLPADHGCSDH 60
Db 383 VSNLGMFGIKDPTAVINPPHATILAVGTGEQRAIVKDGKIEVATWMSVTLSCDHRADVGA 442

QY 61 RCA 63
Db 443 LGA 445

RESULT 11
US-10-893-608-1
; Sequence 1, Application US/10893608
; Publication No. US20050042214A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: Gershwin, Eric M.
; APPLICANT: Leung, Patrick S.C.
; TITLE OF INVENTION: DISCOVERY OF THE MICROORGANISM THAT CAUSES THE HUMAN AUTOIMMUNE
; FILE REFERENCE: UC01190-1
; CURRENT APPLICATION NUMBER: US/10/893,608
; CURRENT FILING DATE: 2004-07-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Novosphingobium aromaticivorans
US-10-893-608-1

Query Match 21.3%; Score 89; DB 17; Length 489;
Best Local Similarity 33.3%; Pred. No. 0.014;
Matches 21; Conservative 11; Mismatches 31; Indels 0; Gaps 0;

QY 1 MTNVGQGVDAVFGVIYPPQVALVSFGKPAQRVCVAVDGAHVMTTVLATLPADHGCSDH 60
Db 409 LSNLGMFSIKNFDAVINPPQAMIMAVGAGEQRYPYVDGALGIATVMSATGDFDHRADGA 468

QY 61 RCA 63

Db 469 DGA 471

RESULT 12
US-10-369-493-21734
; Sequence 21734, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 21734
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Rhodobacter capsulatus
US-10-369-493-21734

Query Match 21.2%; Score 88.5; DB 15; Length 417;
Best Local Similarity 31.4%; Pred. No. 0.014;
Matches 22; Conservative 14; Mismatches 33; Indels 1; Gaps 1;

QY 1 MTNVGQGVDAVFGVIYPPQVALVSFGKPAQR-VCAVDGAHVMTTVLATLPADHGCSDH 59
Db 337 ISNLGMGIESFDVAVINPPHGAIIAVGAGVKKPVVSASGTIEVATMMALTSLSDHVRVIDG 396

QY 60 HRGALFFLSI 69
Db 397 ALGAELLLAAI 406

RESULT 13
US-10-369-493-22915
; Sequence 22915, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 22915
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Aeropyrum pernix
US-10-369-493-22915

Query Match 21.1%; Score 88; DB 15; Length 412;
Best Local Similarity 30.6%; Pred. No. 0.015;
Matches 22; Conservative 10; Mismatches 40; Indels 0; Gaps 0;

QY 1 MTNVGQGVDAVFGVIYPPQVALVSFGKPAQRVCVAVDGAHVMTTVLATLPADHGCSDH 60
Db 326 ITNVGSGSVIGFPPVIYPPNVAIILGVRHLVRPVPYVDGELPKRIGFVSLGFDHRALEGA 385
```

QY 61 RGALFFLSINEL 72
Db 386 YATRFLEVKRL 397

RESULT 14

US-10-369-493-11998
; Sequence 11998, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11998
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Mesorhizobium loti
US-10-369-493-11998

Query Match 21.1%; Score 88; DB 15; Length 454;
Best Local Similarity 27.8%; Pred. No. 0.017;
Matches 20; Conservative 13; Mismatches 39; Indels 0; Gaps 0;
QY 1 MTNVGQGVDAVFGVYPPQVALVSFGKPAQRCAVDGAHVMVTTVLATLPADHGCSDDH 60
Db 374 VSNLWFGIKDFRAVINPPHATILAVGAEERAVVKNGEIKIATVMSVTLSTDHRAVDGA 433
QY 61 RGALFFLSINEL 72
Db 434 LGAELLVAFKRL 445

RESULT 15

US-10-369-493-17795
; Sequence 17795, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17795
; LENGTH: 491
; TYPE: PRT
; ORGANISM: SPHINGOMONAS
US-10-369-493-17795

Query Match 20.8%; Score 87; DB 15; Length 491;
Best Local Similarity 31.9%; Pred. No. 0.026;
Matches 23; Conservative 11; Mismatches 38; Indels 0; Gaps 0;
QY 1 MTNVGQGVDAVFGVYPPQVALVSFGKPAQRCAVDGAHVMVTTVLATLPADHGCSDDH 60
Db 411 LSNMGYGIKQFEAVINPPQGMIMATGAGEKRPYIIDALGVATVMSATGSGFDHRAIDGA 470

QY 61 RGALFFLSINEL 72
Db 471 DGAEMMKIFKEL 482

Search completed: October 15, 2005, 04:15:00
Job time : 19.1311 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2005, 03:59:30 ; Search time 4.46393 Seconds
(without alignments)
1724.342 Million cell updates/sec

Title: US-10-617-038-10

Perfect score: 418

Sequence: 1 MTNVGQGVDAVFGVIYPPQ.....RGALFFLSINELTRCAAVTG 80

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	418	100.0	80	2 A70688	hypothetical prote
2	108	25.8	429	2 B86529	dihydrolipoamide a
3	108	25.8	429	2 B72095	pyruvate dehydroge
4	102	24.4	428	2 F81694	pyruvate dehydroge
5	100	23.9	405	2 AH2752	dihydrolipoamide a
6	100	23.9	457	2 F97533	dihydrolipoamide a
7	99	23.7	429	2 H71539	probable dihydrol
8	97	23.2	428	2 E87463	hypothetical prote
9	91	21.8	404	2 A82121	2-oxoglutarate deh
10	89	21.3	412	2 D97795	hypothetical prote
11	88	21.1	412	2 C72548	probable dihydrol
12	88	21.1	450	2 D36953	dihydrolipoamide s
13	85	20.3	402	2 AE0591	dihydrolipoamide s
14	84	20.1	395	2 B97728	hypothetical prote
15	84	20.1	544	2 C42653	dihydrolipoamide s
16	83.5	20.0	610	2 A55939	dihydrolipoamide s
17	83	19.9	177	2 A99312	hypothetical prote
18	83	19.9	436	2 B83747	dihydrolipoamide s
19	82.5	19.7	637	2 T49099	dihydrolipoamide s
20	82	19.6	401	2 G71728	dihydrolipoamide a
21	82	19.6	405	2 S42875	dihydrolipoamide s
22	82	19.6	409	2 H83448	dihydrolipoamide s
23	81	19.4	402	2 S73773	dihydrolipoamide a
24	80.5	19.3	447	2 AB3359	dihydrolipoamide s
25	80	19.1	399	2 S07779	dihydrolipoamide s
26	80	19.1	405	1 XUECSD	dihydrolipoamide s
27	80	19.1	405	2 G85573	dihydrolipoamide s
28	80	19.1	405	2 H90722	dihydrolipoamide s
29	78.5	18.8	442	2 D36718	dihydrolipoamide s

ALIGNMENTS

RESULT 1

A70688

hypothetical protein Rv1734c - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: A70688

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.,

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:96344230

A:Accession: A70688

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-80 <COL>

A:Cross-references: UNIPROT:P71992; GB:Z81360; GB:AL123456; NID:g3261654; PIDN:CAB03697.1

A:Experimental source: strain H37Rv

C:Genetics:

A:Gene: Rv1734c

Query Match 100.0%; Score 418; DB 2; Length 80;

Best Local Similarity 100.0%; Pred. NO. 2.8e-43;

Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTNVGQGVDAVFGVIYPPQALVSGKPAQRVCAVDGAIHVMTTVLATLPADHGCSDDH 60

DB 1 MTNVGQGVDAVFGVIYPPQALVSGKPAQRVCAVDGAIHVMTTVLATLPADHGCSDDH 60

QY 61 RGALFFLSINELTRCAAVTG 80

DB 61 RGALFFLSINELTRCAAVTG 80

RESULT 2

B86529

dihydrolipoamide acetyltransferase [imported] - Chlamydophila pneumoniae (strain J138)

C:Species: Chlamydophila pneumoniae, Chlamydia pneumoniae

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C:Accession: B86529

R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Ouchi, K.; Shiba, T.; Ise

Nucleic Acids Res. 28, 2311-2314, 2000

A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.

A:Reference number: A86491; MUID:20330349; PMID:10871362

A:Accession: B86529

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-429 <STO>

A:Cross-references: UNIPROT:Q928N2; GB:BA000008; NID:g8978680; PIDN:BAA98516.1; GSPDB:GN

A:Experimental source: strain J138

C:Genetics:

A:Gene: pdhC
C:Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology

Query Match 25.8%; Score 108; DB 2; Length 429;
Best Local Similarity 26.9%; Pred. No. 3.9e-05;
Matches 21; Conservative 19; Mismatches 38; Indels 0; Gaps 0;

QY 1 MTNVGDOGVDVFGVIYPPQVALVSFGKPAQRVCAVDGAIHVMTTVLATLPADHGCSDDH 60
DB 348 VSNLGMTGITEFTAINVPPQAILAVGSVTEQALVLGDIEITGTCNLTSLSVDRHVIDGY 407
QY 61 RGALFFLSINELTRCAAV 78
DB 408 PAAMFMKRLQKLEAPAV 425

RESULT 3
B72095
pyruvate dehydrogenase, E2 component, dihydrolipoamide S-acetyltransferase CP0452 [imported]
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: B72095; A81576
R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammell, C.; Fan, J.; Olinger, L.; Grimwood, J.; Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: B72095
A:Molecule type: DNA
A:Residues: 1-429 <RNA>
A:Cross-references: UNIPROT:Q928N2; GB:AE001615; GB:AE001363; NID:G4376574; PIDN:AAD1845
A:Experimental source: strain CML029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: A81576
A:Molecule type: DNA
A:Residues: 1-429 <REA>
A:Cross-references: GB:AE002206; GB:AE002161; NID:G7189367; PIDN:AAF38290.1; PID:G718937
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: pdhC; CP0452
C:Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology
F:4-77/Domain: lipoyl/biotin-binding homology <LPB>

Query Match 25.8%; Score 108; DB 2; Length 429;
Best Local Similarity 26.9%; Pred. No. 3.9e-05;
Matches 21; Conservative 19; Mismatches 38; Indels 0; Gaps 0;

QY 1 MTNVGDOGVDVFGVIYPPQVALVSFGKPAQRVCAVDGAIHVMTTVLATLPADHGCSDDH 60
DB 348 VSNLGMTGITEFTAINVPPQAILAVGSVTEQALVLGDIEITGTCNLTSLSVDRHVIDGY 407
QY 61 RGALFFLSINELTRCAAV 78
DB 408 PAAMFMKRLQKLEAPAV 425

RESULT 4
F81694
pyruvate dehydrogenase, E2 component, dihydrolipoamide acetyltransferase, probable TC051
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: F81694
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: F81694
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-428 <TET>
A:Cross-references: UNIPROT:Q9PKE7; GB:AE002320; GB:AE002160; NID:G7190551; PIDN:AAF39361
A:Experimental source: strain Nigg (MoPn)
C:Genetics:
A:Gene: TC0518
C:Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology

Query Match 24.4%; Score 102; DB 2; Length 428;
Best Local Similarity 25.6%; Pred. No. 0.00021;
Matches 20; Conservative 21; Mismatches 37; Indels 0; Gaps 0;

QY 1 MTNVGDOGVDVFGVIYPPQVALVSFGKPAQRVCAVDGAIHVMTTVLATLPADHGCSDDH 60
DB 347 ISNLGMTGISDFTAILNPPQAILAVGSVERQPVVLGELAVGTCMLTSLSVDRHVIDGY 406
QY 61 RGALFFLSINELTRCAAV 78
DB 407 PAAMFMKRLQKLEAPSV 424

RESULT 5
AH2752
dihydrolipoamide acetyltransferase [imported] - Agrobacterium tumefaciens (strain C58, Di
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AH2752
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monke, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perty, M.; Gordon-Kamm, I. erer, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AH2752
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-405 <KUR>
A:Cross-references: UNIPROT:Q8UFG5; GB:AE008688; PIDN:AAL42438.1; PID:G17739851; GSPDB:G
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: aceF
A:Map position: circular chromosome

Query Match 23.9%; Score 100; DB 2; Length 405;
Best Local Similarity 33.3%; Pred. No. 0.00034;
Matches 21; Conservative 10; Mismatches 32; Indels 0; Gaps 0;

QY 1 MTNVGDOGVDVFGVIYPPQVALVSFGKPAQRVCAVDGAIHVMTTVLATLPADHGCSDDH 60
DB 325 VSNLGMTGKVSFSAVINPPHATILAVGAGEORAVVKNGEIKIANVMTVLTSDHRCVDGA 384
QY 61 RGA 63
DB 385 LGA 387

RESULT 6
F97533
dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex (e2) [imp
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: F97533
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: F97533
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-457 <KUR>
A:Cross-references: UNIPROT:Q8UFG5; GB:AE007869; PIDN:AAK87223.1; PID:G15156505; GSPDB:G

C;Genetics:
A;Gene: AGR_C_2641
A;Map position: circular chromosome
C;Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology

Query Match 23.9%; Score 100; DB 2; Length 457;
Best Local Similarity 33.3%; Pred. No. 0.00039;
Matches 21; Conservative 10; Mismatches 32; Indels 0; Gaps 0;

QY 1 MTNVGQGVDAVFGVYPPQVALVSGKPAQRCAVDGAIHVMTTVLATLPADHGSDDH 60
Db 377 VSNMGWGVKSFSAVINPHTILAVGAGQRAVVKNGEIKIANVMTVLTSTDRHCV DGA 436

QY 61 RGA 63
Db 437 LGA 439

RESULT 7
H71539
probable dihydrolipoamide acetyltransferase - Chlamydia trachomatis (serotype D, strain C;Species: Chlamydia trachomatis
C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C;Accession: H71539
R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998
A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis
A;Reference number: A71570; MUID:9900809; PMID:9784136
A;Accession: H71539
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-429 <ARN>
A;Cross-references: UNIPROT:O84249; GB:AE001297; GB:AE001273; NID:g3328646; PIDN:AAAC6784
A;Experimental source: serotype D, strain UW-3/Cx
C;Genetics:
A;Gene: pdhC
C;Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology
F;4-77/Domain: lipoyl/biotin-binding homology <LPB1>

Query Match 23.7%; Score 99; DB 2; Length 429;
Best Local Similarity 25.6%; Pred. No. 0.00048;
Matches 20; Conservative 19; Mismatches 39; Indels 0; Gaps 0;

QY 1 MTNVGQGVDAVFGVYPPQVALVSGKPAQRCAVDGAIHVMTTVLATLPADHGSDDH 60
Db 348 VSNLGMTGISDFTAILNPPQAAILAVGSVEQPWVLNGLAVGLTCLMLTSLVDHRVIDGY 407

QY 61 RGAFLFSLINELTRCAAV 78
Db 408 PAAMFMKRLQLLEAPSV 425

RESULT 8
E87463
hypothetical protein CC1729 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: E87463
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: E87463
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-428 <STO>
A;Cross-references: UNIPROT:Q9A7J4; GB:AE005673; NID:g13423147; PIDN:AAK23705.1; GSPDB:C
C;Genetics:
A;Gene: CC1729
C;Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology

Query Match 23.2%; Score 97; DB 2; Length 428;
Best Local Similarity 33.3%; Pred. No. 0.00083;
Matches 24; Conservative 11; Mismatches 37; Indels 0; Gaps 0;

QY 1 MTNVGQGVDAVFGVYPPQVALVSGKPAQRCAVDGAIHVMTTVLATLPADHGSDDH 60
Db 348 ISNLGWFGLKSFASINERPGAIMSVGAGEQRPVVKNGEIKVATVMTVLTCDHRVVDGS 407

QY 61 RGAFLFSLINEL 72
Db 408 VGAKFLAAPRPL 419

RESULT 9
A82121
2-oxoglutarate dehydrogenase, E2 component, dihydrolipoamide succinyltransferase VC2086
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: A82121
R;Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P. 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M. Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: A82121
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-404 <HEI>
A;Cross-references: UNIPROT:Q9KQB4; GB:AE004282; GB:AE003852; NID:g9656626; PIDN:AAF95232;
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC2086
A;Map position: 1
C;Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology

Query Match 21.8%; Score 91; DB 2; Length 404;
Best Local Similarity 30.8%; Pred. No. 0.0041;
Matches 24; Conservative 10; Mismatches 32; Indels 12; Gaps 2;

QY 1 MTNVGQGVDAVFG-----VIYPPQVALVSGKPAQRCAVDGAIHVMTTVLATLPADH 54
Db 322 ITNGG-----VFGSLMSTPIINPPQAAILGHKIQDRAVMVDGKIEILPMYLLALSYDH 375

QY 55 GCSDHHRGALFFLSINEL 72
Db 376 RSIDGRESVGLVTVKEL 393

RESULT 10
D97795
hypothetical protein pdhC [imported] - Rickettsia conorii (strain Malish 7)
C;Species: Rickettsia conorii
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: D97795
R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro, Science 293, 2093-2098, 2001
A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A;Reference number: A97700; MUID:21442074; PMID:11557893
A;Accession: D97795
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-412 <KUR>
A;Cross-references: UNIPROT:Q92HK7; GB:AE006914; PIDN:AALO3302.1; PID:g15619860; GSPDB:G
C;Genetics:
A;Gene: pdhC
C;Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology

Query Match 21.3%; Score 89; DB 2; Length 412;
Best Local Similarity 29.6%; Pred. No. 0.0074;
Matches 21; Conservative 14; Mismatches 36; Indels 0; Gaps 0;

QY 1 MTNVGQGVDAVFGVYPPQVALVSGKPAQRCAVDGAIHVMTTVLATLPADHGSDDH 60

```

Db      332 ISNLGWYGIKFNALINPPQSCIMGVGSARKAIVKYNQIITATIMDTLSADHRVVDGA 391
      61 RGALFFLSINE 71
      392 VQAEPLVAFKK 402

RESULT 11
C72548
probable dihydrolipoamide acetyltransferase component (E2) of pyruvate dehydrogenase com
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: C72548
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: C72548
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-412 <KAW>
A:Cross-references: UNIPROT:Q9YBC6; DDBJ:AP000062; NID:g5105244; PIDN:BAA80672.1; PID:g5
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1671
C:Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology
F:6-79/Domain: lipoyl/biotin-binding homology <LPB>

Query Match      21.1%; Score 88; DB 2; Length 412;
Best Local Similarity 30.6%; Pred. No. 0.0097;
Matches 22; Conservative 10; Mismatches 40; Indels 0; Gaps 0;

QY      1 MTNVGDQGVDAVFGVIYPPQVALVSGFGPAQRVCAVDGAIHVMTTVLATLPADHGCSDDH 60
Db      326 ITNVSGISGVIGFPIYPENVAILGVHRLVRPVYVDGELPKRKIGFVSLGFDHRALEGA 385

QY      61 RGALFFLSINEL 72
Db      386 YATRFLEVKRL 397

RESULT 12
D36953
dihydrolipoamide S-acetyltransferase (EC 2.3.1.12) - Pelobacter carbinolicus
C:Species: Pelobacter carbinolicus
C>Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 09-Jul-2004
C:Accession: D36953
R:Oppermann, F.B.; Steinbuichel, A.
J. Bacteriol. 176, 469-485, 1994
A:Title: Identification and molecular characterization of the acs genes encoding the Pel
A:Reference number: A36953; MUID:94117383; PMID:8110297
A:Accession: D36953
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-450 <OPP>
A:Cross-references: UNIPROT:Q59658; GB:U01100; NTD:g434021; PIDN:AAA18917.1; PID:g434025
C:Genetics:
A:Gene: acoc
C:Keywords: acyltransferase; coenzyme A
F:8-81/Domain: lipoyl/biotin-binding homology <LPB1>
F:123-196/Domain: lipoyl/biotin-binding homology <LPB2>

Query Match      21.1%; Score 88; DB 2; Length 450;
Best Local Similarity 32.9%; Pred. No. 0.011;
Matches 24; Conservative 15; Mismatches 32; Indels 2; Gaps 2;

QY      1 MTNVGDQGVDAVFGVIYPPQV-ALVSGFGPAQRVCAVDGAIHVMTTVLATLPADHGCSDD 59
      372 ISNLGMLGVDS-FGALVPPGSAILAVGGIKDEVVVKDGEVMPVSTMKVTLVADHRVVDG 430
      60 HRGALFFLSINEL 72

```

```

Db      431 LYSAQFLVELKRL 443

RESULT 13
AE0591
dihydrolipoamide succinyltransferase component (E2) [imported] - Salmonella enterica subsp
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A>Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AE0591
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skellon, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AE0591
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-402 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD05199.1; PID:g16501969; GSPDB:GN00176
C:Genetics:
A:Gene: STY0780
C:Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology

Query Match      20.3%; Score 85; DB 2; Length 402;
Best Local Similarity 30.8%; Pred. No. 0.022;
Matches 24; Conservative 10; Mismatches 32; Indels 12; Gaps 2;

QY      1 MTNVGDQGVDAVFG-----VIYPPQVALVSGFGPAQRVCAVDGAIHVMTTVLATLPADH 54
Db      320 ITNGG-----VFGSLMSTPIINPPQSAITGMHAIKRPMAVDGKVELPMWIALSYDH 373

QY      55 GCSDDHRCALFFLSINEL 72
Db      374 RLIDGRESVGLVTIKEL 391

RESULT 14
B97728
hypothetical protein sucB [imported] - Rickettsia conorii (strain Malish 7)
C:Species: Rickettsia conorii
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: B97728
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro
Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: B97728
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-395 <KUR>
A:Cross-references: UNIPROT:Q92J43; GB:AE006914; PIDN:AA02764.1; PID:g15619278; GSPDB:G
C:Genetics:
A:Gene: sucB
C:Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology

Query Match      20.1%; Score 84; DB 2; Length 395;
Best Local Similarity 29.3%; Pred. No. 0.028;
Matches 17; Conservative 9; Mismatches 32; Indels 0; Gaps 0;

QY      15 VIYPPQVALVSGFGPAQRVCAVDGAIHVMTTVLATLPADHGCSDDHRCALFFLSINEL 72
Db      327 IINPPQSGILGHKTEERAVVIDGKIETIRPMWIALSYDHRIDGKGVSLVKIKOL 384

RESULT 15
C42653
dihydrolipoamide S-acetyltransferase (EC 2.3.1.12) - Acholeplasma laidlawii
N:Alternate names: 57K protein
C:Species: Acholeplasma laidlawii

```


Search completed: October 15, 2005, 04:38:28
Job time : 5.46393 secs

THIS PAGE BLANK (USPTO)

Result No.	Score	Query %			DB	ID	Description
		Match	Length				
1	418	100.0	80	2	P71992	mycobacteri	
2	413	98.8	80	2	Q7T2P0	mycobacteri	
3	186	44.5	337	2	Q6ALF2	desulfotale	
4	179	42.8	397	2	Q6A1E3	desulfotale	
5	111.5	26.7	418	2	Q74AE1	geobacter s	
6	108	25.8	428	2	Q9Z8N2	chlamydia p	
7	102	24.4	429	2	Q9PKX7	chlamydia m	
8	100	23.9	405	2	Q8UPG5	agrobacteri	
9	100	23.9	457	2	Q7CZ96	agrobacteri	
10	99	23.7	429	2	Q84249	chlamydia t	
11	97	23.2	428	2	Q823E3	chlamydophi	
12	97	23.2	428	2	Q9A7J4	caulobacter	
13	94	22.5	463	2	Q6NSV6	rhodospseudo	
14	91	21.8	404	2	Q9KQB4	vibrio chol	
15	91	21.8	458	2	Q8F4N2	leptospira	
16	90.5	21.7	442	2	Q6G4O3	bartonella	
17	90	21.5	402	2	Q7WMN4	vibrio vuln	
18	90	21.5	402	2	Q8DFQ0	vibrio vuln	
19	89	21.3	412	1	ODP2_R1CCN	rickettsia	
20	88	21.1	401	2	Q87RE9	vibrio para	
21	88	21.1	412	2	Q9YBC6	aeropyrum p	
22	88	21.1	450	2	Q59E58	peibobacter	
23	88	21.1	454	2	Q98FT5	rhizobium l	
24	88	21.1	470	2	Q8KTE7	methyllobact	
25	87	20.8	412	2	Q7PC39	rickettsia	
26	86	20.6	440	1	ODP2_ZYMMO	zymomonas m	
27	85	20.6	453	2	Q98MY7	rhizobium l	
28	85	20.3	399	2	Q8RBW8	thermoanaer	
29	85	20.3	402	2	Q8XFE6	salmonella	
30	85	20.3	402	2	Q7CQX4	salmonella	
31	84	20.1	395	1	ODO2_R1CCN	rickettsia	

```
DE Hypothetical protein Mbi763c.
GN OrderedLocusNames=Mbi763c;
OS Mycobacterium bovis
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garner T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Gironde S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Bartell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248340; CAD94465.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 80 AA; 8284 MW; A5A9B8216ED9D848 CRC64;

Query Match 98.8%; Score 413; DB 2; Length 80;
Best Local Similarity 98.8%; Pred. No. 3.7e-42;
Matches 79; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTNVGDQGVDAVFGVIYPPQVALVSFGKPAQRVCADVGAHVMVTTLATLPADHGCSDH 60
DB 1 MTNVGDQGVDAVFGVIYPPQVALVSFGKPAQRVCADVGAHVMVTTLATLPADHGCSDH 60

QY 61 RGALFFLSINELTRCAAVTG 80
DB 61 RGALFFLSINELTRCAAVAG 80

RESULT 3
QALF2
ID Q6ALF2 PRELIMINARY; PRT; 397 AA.
AC Q6ALF2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Probable dihydrolipoamide acetyltransferase, component E2 of pyruvate
DE dehydrogenase.
GN Name=phdC; OrderedLocusNames=DP2094;
OS Desulfotalea psychrophila.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacteriales;
OC Desulfobulbaceae; Desulfotalea.
OX NCBI_TaxID=84980;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LSV54 / DSM 12343;
RX PubMed=15305914;
RA Rabus R., Ruepp A., Frickey T., Rattai T., Fartmann B., Stark M.,
RA Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K.,
RA Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,
RA Klenk H.-P.;
RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
RT from permanently cold Arctic sediments.";
RL Environ. Microbiol. 6:887-902(2004).
CC -!- SIMILARITY: Belongs to the 2-oxoacid dehydrogenase family.
CC -!- SIMILARITY: Contains 1 lipoyl-binding domain.
DR EMBL; CR522871; CAG37904.1; -.
DR GO; GO:0008415; F:acetyltransferase activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001078; 2-oxoacid dh.
DR InterPro; IPR000089; Biotin lipoyl.
DR InterPro; IPR004167; E3 binding.
DR Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,
RA Klenk H.-P.;
RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
RT from permanently cold Arctic sediments.";
RL Environ. Microbiol. 6:887-902(2004).
CC -!- SIMILARITY: Belongs to the 2-oxoacid dehydrogenase family.
CC -!- SIMILARITY: Contains 1 lipoyl-binding domain.
DR EMBL; CR522870; CAG36823.1; -.
DR GO; GO:0008415; F:acetyltransferase activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001078; 2-oxoacid dh.
DR InterPro; IPR000089; Biotin lipoyl.
DR InterPro; IPR004167; E3 binding.
DR InterPro; IPR011053; Hybrid motif.
DR InterPro; IPR003016; Lipoyl_BS.
DR Pfam; PF00198; 2-oxoacid dh; 1.
DR Pfam; PF00364; Biotin lipoyl; 1.
DR Pfam; PF02817; E3_binding; 1.

Query Match 42.8%; Score 179; DB 2; Length 397;
Best Local Similarity 50.0%; Pred. No. 4.7e-13;
Matches 35; Conservative 13; Mismatches 22; Indels 0; Gaps 0;

QY 1 MTNVGDQGVDAVFGVIYPPQVALVSFGKPAQRVCADVGAHVMVTTLATLPADHGCSDH 60
DB 319 ITSGLDGLVEVHGVIYPPQVALVSFGKPAQRVCADVGAHVMVTTLATLPADHGCSDH 378
QY 61 RGALFFLSIN 70
DB 61 RGALFFLSIN 70

RESULT 4
Q6AIE3
ID Q6AIE3 PRELIMINARY; PRT; 397 AA.
AC Q6AIE3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Probable pyruvate dehydrogenase, E2 component, dihydrolipoamide
DE acetyltransferase.
GN OrderedLocusNames=DPB40;
OS Desulfotalea psychrophila.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacteriales;
OC Desulfobulbaceae; Desulfotalea.
OX NCBI_TaxID=84980;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LSV54 / DSM 12343;
RX PubMed=15305914;
RA Rabus R., Ruepp A., Frickey T., Rattai T., Fartmann B., Stark M.,
RA Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K.,
RA Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,
RA Klenk H.-P.;
RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
RT from permanently cold Arctic sediments.";
RL Environ. Microbiol. 6:887-902(2004).
CC -!- SIMILARITY: Belongs to the 2-oxoacid dehydrogenase family.
CC -!- SIMILARITY: Contains 1 lipoyl-binding domain.
DR EMBL; CR522871; CAG37904.1; -.
DR GO; GO:0008415; F:acetyltransferase activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001078; 2-oxoacid dh.
DR InterPro; IPR000089; Biotin lipoyl.
DR InterPro; IPR004167; E3 binding.
DR InterPro; IPR011053; Hybrid motif.
DR InterPro; IPR003016; Lipoyl_BS.
DR Pfam; PF00198; 2-oxoacid dh; 1.
DR Pfam; PF00364; Biotin lipoyl; 1.
DR Pfam; PF02817; E3_binding; 1.
DR ProDom; PD001115; 2-oxoacid dh; 1.
DR PROSITE; PS00189; LIPOYL; 1.
KW Acyltransferase; Complete proteome; Lipoyl; Pyruvate; Transferase.
SQ SEQUENCE 397 AA; 42829 MW; 528EB659FAC32D9B CRC64;

Query Match 52.9%; Score 186; DB 2; Length 397;
Best Local Similarity 52.9%; Pred. No. 6.6e-14;
Matches 37; Conservative 11; Mismatches 22; Indels 0; Gaps 0;

QY 1 MTNVGDQGVDAVFGVIYPPQVALVSFGKPAQRVCADVGAHVMVTTLATLPADHGCSDH 60
DB 319 ITNLGDLGVEVHGVIYPPQVALVSFGKPAQRVCADVGAHVMVTTLATLPADHGCSDH 378
QY 61 RGALFFLSIN 70
DB 379 RGAQFLEALN 388

Query Match 44.5%; Score 186; DB 2; Length 397;
Best Local Similarity 44.5%; Pred. No. 6.6e-14;
Matches 37; Conservative 11; Mismatches 22; Indels 0; Gaps 0;
```

Name=pbdc; OrderedLocusNames=CP0452, CP0306, CPB0315;
 Chlamydia pneumoniae (Chlamydiales: Chlamydiaceae).
 Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.

NCBI_TaxID=83558;
 [1]
 SEQUENCE FROM N.A.
 RP STRAIN=CWL029;
 RC MEDLINE=9920606; PubMed=10192388; DOI=10.1038/7716;
 RX Kalman S., Mitchell W.P., Marathe R.W., Lammel C.J., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R., Stephens R.S.;
 RA "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";

Nat. Genet. 21:385-389(1999).
 [2]
 RP "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
 RT pneumoniae AR39.";

RC Nucleic Acids Res. 28:1397-1406(2000).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AR39;
 RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,
 RA Badd S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
 RA Dosson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
 RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
 RT pneumoniae AR39.";

RC Nucleic Acids Res. 28:1397-1406(2000).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TW-183;
 RX Gang M.M., Schuhmacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,
 RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
 RA "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
 RT other Chlamydia strains based on whole genome sequence analysis";
 RT Submitted (MAY-2002) to the EMBL/Genbank/DBSJ databases.
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J138;
 RX MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
 RA "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RT from Japan and CWL029 from USA.";

RC Nucleic Acids Res. 28:2311-2314(2000).
 CC -1- SIMILARITY: Belongs to the 2-oxoacid dehydrogenase family.
 DR ENBL; AE001615; AAD18455.1; -;
 DR ENBL; AE002205; AAF38290.1; -;
 DR ENBL; AE01158; AAP98248.1; -;
 DR ENBL; AP002546; BAA98516.1; -;
 DR PIR; B72095; B72095;
 DR PIR; B86529; B86529;
 DR HSSP; P07016; 1BBL;
 DR PHC1-2DPAGE; Q9Z8N2; -;
 DR TIGR; CP0452; -;
 DR GO; GO:0045254; C:pyruvate dehydrogenase complex; IEA.
 DR GO; GO:0009415; F:acyltransferase activity; IEA.
 DR GO; GO:0004742; F:dihydrolypyllysine-residue acetyltransferase.; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR001078; 2oxoacid_dh.
 DR InterPro; IPR006257; AcettransferaseL.
 DR InterPro; IPR00089; Biotin lipoyl.
 DR InterPro; IPR004167; E3 binding.
 DR InterPro; IPR011053; HyBrid motif.
 DR Pfam; PF00198; 2-oxoacid_dh; 1.
 DR Pfam; PF00364; Biotin lipoyl; 1.
 DR Pfam; PF02817; E3 binding; 1.
 DR ProDom; PD001115; 2oxoacid_dh; 1.
 DR TIGRFAMS; TIGR01349; pDHac trf_mito; 1.
 KW Acyltransferase; Complete Proteome; Lipoyl; Pyruvate; Transferase.
 SQ SEQUENCE 429 AA; 46256 MW; B2BD3334FCB493C5 CRC64;

Query Match 25.8%; Score 109; DB 2; Length 429;
 Best Local Similarity 26.9%; Pred No. 0.00021;

```

Matches 21; Conservative 19; Mismatches 38; Indels 0; Gaps 0;
QY 1 MTNVGQGVDAVFGVIYPPQVALVSGKPAQRVCAGVCAVDAIHWMTTVLATLPADHGCSDDH 60
Db VSNLGMTGITEFTAINPFPQAIALVSGVTEQALVLDGEITGSTCNLTLSVDHRVIDGY 407
QY 61 RGALFFLSINELTRCAAV 78
Db PAAMFMKRLQKLEAPAV 425

RESULT 7
Q9PK7 PRELIMINARY; PRT; 428 AA.
AC Q9PK7;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Pyruvate dehydrogenase, E2 component, dihydrolipoamide
DE acetyltransferase, putative.
CN OrderedLocusNames=TC0518;
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MoPn / Ni99;
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,
RA Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
RA Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.,
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -!- SIMILARITY: Belongs to the 2-oxoacid dehydrogenase family.
DR EMBL; AE002320; AAF39360.1; -.
DR PIR; F81694; F81694.
DR HSSP; P07016; 1BBL.
DR TIGR; TC0518; -.
DR GO; GO:0045254; C:pyruvate dehydrogenase complex; IEA.
DR GO; GO:0008415; F:acetyltransferase activity; IEA.
DR GO; GO:0004742; F:dihydrolipoyllysine-residue acetyltransferase. . .; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001078; 2Oxoacid dh.
DR InterPro; IPR006257; AceTtransferaseL.
DR InterPro; IPR000089; Biotin lipoyl.
DR InterPro; IPR004167; E3_binding.
DR InterPro; IPR011053; Hybrid_motif.
DR Pfam; PF001198; 2-oxoacid dh; 1.
DR Pfam; PF00364; Biotin lipoyl; 1.
DR Pfam; PF02817; E3 binding; 1.
DR ProDom; PD001115; 2Oxoacid dh; 1.
DR TIGRfams; TIGR01349; PDHac_trf_mito; 1.
KW Acyltransferase; Complete proteome; Lipoyl; Pyruvate; Transferase.
SQ SEQUENCE 428 AA; 46177 MW; 4AB96BE4AF01E9FE CRC64;

Query Match 24.4%; Score 102; DB 2; Length 428;
Best Local Similarity 25.6%; Pred. No. 0.0011;
Matches 20; Conservative 21; Mismatches 37; Indels 0; Gaps 0;
QY 1 MTNVGQGVDAVFGVIYPPQVALVSGKPAQRVCAGVCAVDAIHWMTTVLATLPADHGCSDDH 60
Db ISNLGMTGISDFTAINPFPQAIALVSGVEEQPVVLNGELAVGSTCMLTSLVDHRVIDGY 406
QY 61 RGALFFLSINELTRCAAV 78
Db PAAMFMKRLQKLEAPSV 424

Matches 21; Conservative 19; Mismatches 38; Indels 0; Gaps 0;
QY 1 MTNVGQGVDAVFGVIYPPQVALVSGKPAQRVCAGVCAVDAIHWMTTVLATLPADHGCSDDH 60
Db VSNLGMTGITEFTAINPFPQAIALVSGVTEQALVLDGEITGSTCNLTLSVDHRVIDGY 407
QY 61 RGALFFLSINELTRCAAV 78
Db PAAMFMKRLQKLEAPSV 425

RESULT 8
Q9UG5 PRELIMINARY; PRT; 405 AA.
AC Q9UG5;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Dihydrolipoamide acetyltransferase.
GN Name=aceF; OrderedLocusNames=Atu1432;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dupont;
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. St., Woo L.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kuyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
CC -!- SIMILARITY: Belongs to the 2-oxoacid dehydrogenase family.
DR EMBL; AE009103; AAL42438.1; -.
DR PIR; AH2752; AH2752.
DR PIR; F97533; F97533.
DR HSSP; P07016; 1BAL.
DR GO; GO:0045254; C:pyruvate dehydrogenase complex; IEA.
DR GO; GO:0008415; F:acetyltransferase activity; IEA.
DR GO; GO:0004742; F:dihydrolipoyllysine-residue acetyltransferase. . .; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR Pfam; PF001198; 2-oxoacid dh; 1.
DR ProDom; PD001115; 2Oxoacid dh; 1.
DR TIGRfams; TIGR01349; PDHac_trf_mito; 1.
KW Complete proteome; Transferase.
SQ SEQUENCE 405 AA; 41861 MW; B6FF49ACA639F58C CRC64;

Query Match 23.9%; Score 100; DB 2; Length 405;
Best Local Similarity 33.3%; Pred. No. 0.0018;
Matches 21; Conservative 10; Mismatches 32; Indels 0; Gaps 0;
QY 1 MTNVGQGVDAVFGVIYPPQVALVSGKPAQRVCAGVCAVDAIHWMTTVLATLPADHGCSDDH 60
Db VSNMGMGVKSFSVAVINPPHATILAVGAEQRAVVKNGEIKIANVMTVTLSTDHRCVDA 384
QY 61 RGA 63
Db 385 LGA 387

RESULT 9
Q7CZ96 PRELIMINARY; PRT; 457 AA.
AC Q7CZ96;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE AGR_C_2641p.
GN OrderedLocusNames=AGR_C_2641;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

```

```

OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=Cereon;
RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmel K., Gordon J., Vaudin M., Iatchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
CC -!- SIMILARITY: Belongs to the 2-oxoacid dehydrogenase family.
CC -!- SIMILARITY: Contains 1 lipoyl-binding domain.
DR EMBL; AE008068; AA87223.1; -.
DR HSSP; P07016; 1BAL..
DR GO; GO:0045254; C:pyruvate dehydrogenase complex; IEA.
DR GO; GO:0008415; F:acetyltransferase activity; IEA.
DR GO; GO:0004742; F:dihydrolipoylllysine-residue acetyltransferase.; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006096; F:glycolysis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR DR InterPro; IPR001078; 2Oxoacid dh.
DR DR InterPro; IPR006257; AceFtransferasel.
DR DR InterPro; IPR000089; Biotin_lipoyl.
DR DR InterPro; IPR011053; Hybrid_motif.
DR DR InterPro; IPR003016; Lipoyl_BS.
DR Pfam; PF00198; 2-oxoacid dh; 1.
DR Pfam; PF00364; Biotin_lipoyl; 1.
DR ProDom; PD001115; 2Oxoacid dh; 1.
DR TIGRfams; tigr01349; PDHac_trf_mito; 1.
DR PROSITE; PS00189; LIPOYL; 1.
KW Acyltransferase; Lipoyl; Transferase.
SQ SSSEQUENCE 457 AA; 47384 MW; 24BC58D036E71976 CRC64;

Query Match 23.9%; Score 100; DB 2; Length 457;
Best Local Similarity 33.3%; Pred.No. 0.002;
Matches 21; Conservative 10; Mismatches 32; Indels 0; Gaps 0

QY 1 MTNVGQGVDAVFGVIYPQVALVSFGKPAQRCAVDGAIHVMVTVLATLPADHGCSDDH 630
Db 377 VS:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 61 RGA 63
Db 437 LGA 439

RESULT 10
OB4249 PRELIMINARY; PRT; 429 AA.
AC OB4249;
CD 01-NOV-1998 (TrEMBLrel_08, Created)
DT 01-NOV-1998 (TrEMBLrel_08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel_25, Last annotation update)
DE Dihydroilpoamide Acetyltransferase.
GN Names=phdC; OrderedLocusNames=CT247;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=D / UW-3 / Cx;
RC MEDLINE=99000809; PubMed=9784136; DOI=10.1126/science.282.5389.754;
RA Stephens R.S., Kaiman S., Kamel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatunov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
```

Science 282:754-759(1998).

-I- SIMILARITY: Belongs to the 2-oxoacid dehydrogenase family.

EMBL: AE001297; AAC67840.1; -

PIR: H71539; H71539.

HSP: P07016; 1BAL.

GO: GO:0045254; C:pyruvate dehydrogenase complex; IEA.

GO: GO:0008415; F:acyltransferase activity; IEA.

GO: GO:0004742; F:diacylglycerollysine-residue acetyltransferase.; IEA.

GO: GO:0005515; F:protein binding; IEA.

GO: GO:0016740; F:transferase activity; IEA.

GO: GO:0006096; P:glycolysis; IEA.

GO: GO:0008152; P:metabolism; IEA.

InterPro: IPR001078; 2oxoacid_dh.

InterPro: IPR006257; AcetyltransferaseL.

InterPro: IPR000089; Biotin lipoyl.

InterPro: IPR004167; E3 binding.

InterPro: IPR011053; Hybrid motif.

Pfam: PF00198; 2-oxoacid_dh; 1.

Pfam: PF02364; Biotin_lipoyl; 1.

Pfam: PF02817; E3 binding; 1.

ProDom: PD001115; 2oxoacid dh; 1.

TIGRFAMs: TIGR01349; PDHAc trf_mito; 1.

Acyltransferase; Complete proteome; Lipoyl; Transferase.

SEQUENCE 429 AA; 46321 MW; 1D73D66317F1180 CRC64;

Query Match 23.7%; Score 99; DB 2; Length 429;

Best Local Similarity 25.6%; Pred. No. 0.0025;

Matches 20; Conservative 19; Mismatches 39; Indels 0; Gaps 0;

Qy 1 MTNVGQGVDAVFGVTPPOVALVSGKPAQRCAVGDGAIHVMTVLATLPADHGSGDDH 60

Db 348 VSNLGTIGTDFAILNPPQAAIILAVSGVEQPVNLNGELAVGLTCLTSLVDHRVDIGY 407

Qy 61 RGALFFLSINELTRCAAV 78

Db 408 PAAMFKRLQRLLEAPSV 425

RESULT 11

Q823E3 PRELIMINARY; PRT; 428 AA.

ID Q823E3

AC Q823E3

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Pyruvate dehydrogenase, E2 component, diacylglycerolamide S-acetyltransferase.

DE acetyltransferase.

GN Names:pdhc; OrderedLocusNames=CCA00476;

OS Chlamydomophila caviae.

OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.

OX NCBI_TaxID=83557;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=GPIC;

RX MEDLINE=325569155; PubMed=12682364; DOI=10.1093/nar/gkg321;

RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T., Heidelberg J.F., Holtzaple E.K., Khouri H.M., Federova N.B., Carty H.A., Umayam L.A., Haft D.-H., Peterson J.D., Beanan M.J., White O., Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G., Ravoi P.M., Fraser C.M.;

RT "Genome sequence of Chlamydomophila caviae (Chlamydia psittaci GPIC): examining the role of niche-specific genes in the evolution of the Chlamydiae".

RL Nucleic Acids Res. 31:2134-2147(2003).

CC -I- SIMILARITY: Belongs to the 2-oxoacid dehydrogenase family.

EMBL: AE016995; AAP05221.1; -

DR HSP: P07016; 1BBL.

DR TIGR: CCA00476; -

DR GO: GO:0045254; C:pyruvate dehydrogenase complex; IEA.

DR GO: GO:0008415; F:acyltransferase activity; IEA.

DR GO: GO:0004742; F:diacylglycerollysine-residue acetyltransferase.; IEA.

DR GO: GO:0005515; F:protein binding; IEA.

DR GO: GO:0016740; F:transferase activity; IEA.

DR	GO: GO:0008096; P:glycolysis; IEA.	
DR	GO: GO:0008152; P:metabolism; IEA.	
DR	InterPro: IPR001078; 2Oxoacid dh.	
DR	InterPro: IPR006257; AcetyltransferaseL.	
DR	InterPro: IPR000089; Biotin lipoyl.	
DR	InterPro: IPR004167; E3 binding.	
DR	InterPro: IPR01053; Hybrid_motif.	
DR	Pfam: PF00198; 2-oxoacid dh; 1.	
DR	Pfam: PF00364; Biotin lipoyl; 1.	
DR	Pfam: PF02817; E3 binding; 1.	
DR	ProDom: PD001115; 2Oxoacid dh; 1.	
DR	TIGRFAMs: TIGR01349; PDHac_Trf_mito; 1.	
KW	Acyltransferase; Complete hydrogenase; Lipoyl; Pyruvate; Transferase.	
SEQ	SEQUENCE 428 AA; 46319 MW; 02037A7266E0D4B0 CRC64;	
	Query Match 23.2%; Score 97; DB 2; Length 428;	
	Best Local Similarity 25.6%; Pred No. 0.0044;	
	Matches 20; Conservative 19; Mismatches 39; Indels 0; Gaps 0	
QY	1 MTNVGDSQGVDAVGVYIPQVALVSFGKPAQRCAVGDGAIHVMVTTLATLPADHGCSDDH 60	
DB	347 VSNLGTGTGITEFTAIINPQAAILAVGSQVEPVMNGEIVAGSTCILTILSIDHRVIDGY 406	
QY	61 RGALFSLINELTRCAAV 78	
DB	407 PAAMFMKRLQKILEAPSV 424	
RESULT 12		
Q9AJV4		
ID	Q9AJV4 PRELIMINARY; PRT; 428 AA.	
AC	Q9AJV4;	
DT	01-JUN-2001 (TrEMBLrel. 17, Created)	
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)	
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	
DE	Pyruvate dehydrogenase complex, E2 component, dihydrolipoamide acetyltransferase.	
GN	OrderedLocusNames=CC1729;	
OS	Caulobacter crescentus.	
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;	
OC	Caulobacteraceae; Caulobacter.	
OX	NCBI_TaxID=155892;	
RP	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=ATCC 19089 / CB15;	
RC	MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;	
RA	Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,	
RA	Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,	
RA	Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,	
RA	DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,	
RA	Kolony J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,	
RA	Berry K.J., Uterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,	
RA	Ermoelova M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,	
RA	Fraser C.M.;	
RT	"Complete genome sequence of Caulobacter crescentus.";	
RT	Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).	
CC	-1- SIMILARITY: Belongs to the 2-oxoacid dehydrogenase family.	
CC	-1- SIMILARITY: Contains 1 lipoyl-binding domain.	
DR	ENBL; AE003847; AAK23705.1; -.	
DR	PIR; E87463; E87463.	
DR	HSSP; P07016; 1BLB.	
DR	TIGR; CC1729; -.	
DR	GO: GO:0045254; C:pyruvate dehydrogenase complex; IEA.	
DR	GO: GO:0008415; Fracyltransferase activity; IEA.	
DR	GO: GO:0004742; F:diacylpolypyllysine-residue acetyltransferase. .; IEA.	
DR	GO: GO:0005515; F:protein binding; IEA.	
DR	GO: GO:0016740; F:transferase activity; IEA.	
DR	GO: GO:0006096; P:glycolysis; IEA.	
DR	GO: GO:0008152; P:metabolism; IEA.	
DR	InterPro: IPR001078; 2Oxoacid dh.	
DR	InterPro: IPR006257; AcetyltransferaseL.	
DR	InterPro: IPR000089; Biotin lipoyl.	
DR	InterPro: IPR004167; E3 binding.	

Best Local Similarity 34.9%; Pred. No. 0.011;
Matches 22; Conservative 8; Mismatches 33; Indels 0; Gaps 0;

QY 1 MTNVGQGVDAVFGVIYPPQVALVSFGKPAQRCAVDGAHVMVTTLATLPADHGCSDDH 60
Db 383 VSNLGMFGIKDFTAVINPHTATILAVGTGQRAIVKDGKIEVATMVSVTLSCHDRAVDGA 442

QY 61 RGA 63

Db 443 LGA 445

RESULT 14

Q9KQB4 ID Q9KQB4 PRELIMINARY; PRT; 404 AA.

AC Q9KQB4; 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE 2-oxoglutarate dehydrogenase, E2 component, dihydroliipoamide
DE succinyltransferase.
GN OrderedLocusNames=VC2086;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,
RA Ermolaeva M.D., Vamathevan J.J., Bass S., Qin H., Dragoi I.,
RA Sellers P., McDonald L.A., Uterback T.R., Fleischmann R.D.,
RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
RT *cholerae*."

RL Nature 406:477-483(2000).
CC -!- SIMILARITY: Belongs to the 2-oxoacid dehydrogenase family.
CC -!- SIMILARITY: Contains 1 lipoyl-binding domain.
DR EMBL; AE004282; AAF9532.1; -;
DR HSSP; P07016; 1E20.
DR TIGR; VC2086; -;
DR GO; GO:0045252; C:oxoglutarate dehydrogenase complex; IEA.
DR GO; GO:0008415; F:acyltransferase activity; IEA.
DR GO; GO:0004149; F:dihydrolipoyllysine-residue succinyltransfe. .; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0006099; P:tricarboxylic acid cycle; IEA.
DR InterPro; IPR001078; 2Oxoacid dh.
DR InterPro; IPR000089; Biotin lipoyl.
DR InterPro; IPR004167; E3 binding.
DR InterPro; IPR011053; Hybrid motif.
DR InterPro; IPR003016; Lipoyl_BS.
DR InterPro; IPR006255; SucB.
DR Pfam; PF00198; 2-oxoacid dh; 1.
DR Pfam; PF00364; Biotin lipoyl; 1.
DR Pfam; PF02817; E3 binding; 1.
DR ProDom; PD001115; 2Oxoacid dh; 1.
DR TIGRFAMs; TIGR01347; sucB; 1.
DR PROSITE; PS00189; LIPOYL; 1.

KW Acyltransferase; Complete proteome; Lipoyl; Transferase.
SQ SEQUENCE 404 AA; 44090 MW; CAC553F82ED5A2AD CRC64;
Query Match 21.8%; Score 91; DB 2; Length 404;
Best Local Similarity 30.8%; Pred. No. 0.022;
Matches 24; Conservative 10; Mismatches 32; Indels 12; Gaps 2;

QY 1 MTNVGQGVDAVFG-----VIYPPQVALVSFGKPAQRCAVDGAHVMVTTLATLPADH 54

Db 322 ITNGG-----VFGSLMSTPIINPPQAILGMHKIQDRAWVDGKIEILPMMYLSYDH 375

QY 55 GCSDDHRGALFELSINEL 72

Db 376 RSDIGRESVGLVTVKEL 393

RESULT 15

Q8F4N2 ID Q8F4N2 PRELIMINARY; PRT; 458 AA.

AC Q8F4N2; 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase
DE complex E2 (EC 2.3.1.12).
GN Name=sdhC; OrderedLocusNames=LA2008;
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX MEDLINE=22598143; PubMed=12712204; DOI=10.1038/nature01597;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
RA Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
RA Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia O.-C., Guo X.-K., Danchin A.,
RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
RA Xu J.-G., Zhao G.-P.;
RT "Unique physiological and pathogenic features of *Leptospira*
RT *interrogans* revealed by whole-genome sequencing."

RL Nature 422:888-893(2003).
CC -!- SIMILARITY: Belongs to the 2-oxoacid dehydrogenase family.
CC -!- SIMILARITY: Contains 1 lipoyl-binding domain.
DR EMBL; AB011373; AAN49207.1; -;
DR HSSP; P10515; 1FYC.
DR GO; GO:0008415; F:acyltransferase activity; IEA.
DR GO; GO:0004742; F:dihydrolipoyllysine-residue acetyltransferase. .; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001078; 2Oxoacid dh.
DR InterPro; IPR000089; Biotin lipoyl.
DR InterPro; IPR004167; E3 binding.
DR InterPro; IPR011053; Hybrid motif.
DR InterPro; IPR003016; Lipoyl_BS.
DR Pfam; PF00198; 2-oxoacid dh; 1.
DR Pfam; PF00364; Biotin lipoyl; 1.
DR Pfam; PF02817; E3 binding; 1.
DR ProDom; PD001115; 2Oxoacid dh; 1.
DR PROSITE; PS00189; LIPOYL; 1.

KW Acyltransferase; Complete proteome; Lipoyl; Transferase.
SQ SEQUENCE 458 AA; 49230 MW; 51F9C09A6F8EFF0F CRC64;
Query Match 21.8%; Score 91; DB 2; Length 458;
Best Local Similarity 30.1%; Pred. No. 0.025;
Matches 22; Conservative 15; Mismatches 36; Indels 0; Gaps 0;

QY 1 MTNVGQGVDAVFGVIYPPQVALVSFGKPAQRCAVDGAHVMVTTLATLPADHGCSDDH 60

Db 377 VSNLGMFGISSFTAVINEPEAAILAVGALVEKPSIVVGKTLNLTLSCHDHRVVDGA 436

QY 61 RGALFELSINELT 73

Db 437 TGARFLSSPRDYT 449

Search completed: October 15, 2005, 04:35:59
Job time : 18.9195 secs

THIS PAGE BLANK (USPTO)